Capturing Protein Dynamics with Time-Resolved Luminescence Spectroscopy

Thesis by

Nicole Bouley Ford

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2013

(Defended May 1, 2013)

Abstract

The presented doctoral research utilizes time-resolved spectroscopy to characterize protein dynamics and folding mechanisms. We resolve millisecond-timescale folding by coupling time-resolved fluorescence energy transfer (trFRET) to a continuous flow microfluidic mixer to obtain intramolecular distance distributions throughout the folding process. We have elucidated the folding mechanisms of two cytochromes—one that exhibits two-state folding (cytochrome cb_{562}) and one that has both a kinetic refolding intermediate ensemble and a distinct equilibrium unfolding intermediate (cytochrome c_{552}). Our data reveal that the distinct structural features of cytochrome c_{552} contribute to its thermostability.

We have also investigated intrachain contact dynamics in unfolded cytochrome cb_{562} by monitoring electron transfer, which occurs as the heme collides with a ruthenium photosensitizer, covalently bound to residues along the polypeptide. Intrachain diffusion for chemically denatured proteins proceeds on the microsecond timescale with an upper limit of 0.1 μs. The power-law dependence (slope = -1.5) of the rate constants on the number of peptide bonds between the heme and Ru complex indicate that cytochrome cb_{562} is minimally frustrated.

In addition, we have explored the pathway dependence of electron tunneling rates between metal sites in proteins. Our research group has converted cytochrome b_{562} to a *c*-type cytochrome with the porphyrin covalently bound to cysteine sidechains. We have investigated the effects of the changes to the protein structure (i.e., increased rigidity and potential new equatorial tunneling pathways) on the electron transfer rates, measured by transient absorption, in a series of ruthenium photosensitizer-modified proteins.

Acknowledgements

When I think back on my time in graduate school, I remember how my friends, coworkers, and family have been there to support me when times were tough. So I wanted to say "thanks" to anyone and everyone who has ever been a part of my support network.

First and foremost, I could not have asked for a more supportive advisor than Harry Gray. My graduate school experience has been defined by being a member of the Gray Nation. Harry keeps an open door and a smile on his face. He cares about the people in his group, not just results. Thanks also, Harry, for giving me the opportunity to go to conferences and to Japan. I appreciate your mentoring during group meetings, while writing papers, and in professional matters.

Jay Winkler has been an excellent advisor in all things technical, from lasers to paper-writing. Discussions with Jay about my results have influenced how I think about biophysical problems. Thanks, Jay, for welcoming me into the group when I became interested in the protein folding projects.

I have been fortunate to have all of my committee members at my candidacy exam, fourth-year meeting, proposal exam, and thesis defense. Jackie Barton, my committee chair, has given me focus and helpful advice. Tom Miller has given

my research and proposals considerable thought and always has an insightful, complex physical question to ask. Douglas Rees has been helpful in discussions of protein dynamics.

My work with Sony has really broadened my graduate school experience. Seiji Yamada is a good friend of mine with whom I enjoyed collaborating. Thank you, Seiji, for inviting me to experience Japan with you and your family. I would like to thank Tokita-san for supporting me while I worked in his group at Sony. Discussions with Goto-san about science and culture contributed to my experience at Sony. I also appreciate the rest of the biomaterial research group who welcomed me into their group. Arigatou gozaimasu. I would also like to acknowledge my collaborator on the electron tunneling project, Tatiana Prytkova.

I have been fortunate to work with the wonderful members of the Gray Group, and especially closely with the bio subgroup: Bert, Tetsu, Melanie, Heather, Keiko, Crystal, Gretchen, Charlotte, Kyle, Maraia, Lionel, Matt, Jeff, Astrid, Peter, Oliver, and Kana. We had great times, in and out of lab. Bert recruited me into the group and taught me how to get around lab, express proteins, and think about life after graduate school. Tetsu taught me how to use the microfluidic mixer. Melanie has been a good friend, exercise buddy, and fellow biotech enthusiast. Heather was a great roommate and is my favorite person to talk to about cats, then science, and then cats. Gretchen was my partner in learning the intricacies of the picosecond laser system, and Kana has been a fun laser buddy. I appreciate the feedback while writing my proposals from my labmates: Peter, Maraia, Kana, Melanie, Heather,

Bert, and Gretchen. I would also like to acknowledge the hard work of the two undergraduate researchers whom I have mentored—Dong Woo Shin and Yuehan Huang.

My friends outside of lab have been a great support group for me. Andrea, Alma, Eldar, and Artur were great fun and helped me to relax at the end of the week. I love catching up with Lisa, Caitlin, Chethana, Amanda, Leslie, and Linhda (my link to Los Angeles). The Miller group has always welcomed me into their fun troupe: Artur, Nick, Jason, Josh, and Connie. My fellow classmates, the "first years," helped in the transition to Caltech—to Long, you are missed. Involvement with the Caltech Y has been a source of great enjoyment and a productive break from work; I would like to thank the Y employees (Greg, Athena, Liz), my fellow Y Outdoor Committee members, especially Deva and Avni, and the students who planned cultural and service events.

Many people in my past have believed in me and helped me make it to Caltech. My mom has always been supportive of me and shows her love through the little things, like letters, homemade pie, and chocolate. My dad has shown me how important it is to keep learning throughout life and was influential in my decision to go to college at Trinity, where there were many opportunities for me. My sisters, Regina and Daphne, were always excited about me doing cool science. They have grown up from "PIGs" to be smart, beautiful ladies. My brother-in-law David has been a wonderful addition to the family. My friends Elizabeth and Rebecca "dreamed big" with me during the difficult years of high school. My college

roommates Lauren, Christie, and Sarah are amazing people and my closest friends. Trinity University chemistry professors (Drs. Urbach, Bachrach, Pursell, and Mills) are great teachers and scientists who inspired me to go to graduate school.

And most importantly, I do not know where I would be without Will, my love and partner who believes in me. He helped me get through candidacy and has shared with me his knowledge of kinetics, data fitting, and biology. He has emotionally supported me throughout. Life is more enjoyable with Will, and Caltech would not have been what it is for me without him.

Contents

List of Figures

List of Tables

Chapter 1

Background

1.1 Our Approach

We are interested in dynamic processes in proteins that occur over a broad range of timescales: protein folding, intrachain diffusion in unfolded proteins, and electron tunneling through proteins. Small-scale motions such as side-chain rotations in proteins occur in picoseconds, whereas large-scale motions such as intrachain diffusion occur in microseconds. The entire folding reaction for a protein can take microseconds to seconds, depending on the protein.

We probe these dynamics by attaching a fluorescent photosensitizer to the protein. We investigate the process on a suitable laser system, which must excite the sample significantly faster than the timescale of the process. For example, we monitor electron transfer events that occur on a microsecond timescale with a nanosecond laser and a long-lived (microseconds) luminescent ruthenium complex. Conformational rearrangements in unfolded proteins, which occur on a similar timescale, are also studied with this system.

We measure intramolecular distances in proteins using a picosecond laser

system and an organic fluorophore with an excited state lifetime of tens of nanoseconds. Since large-scale motions occur significantly slower than the nanosecond regime, we capture a ”snapshot” of the ensemble of protein conformations. We can follow the folding process by collecting snapshots of protein ensembles over milliseconds; this is carried out by coupling the picosecond laser system to a continuous flow mixer to trigger the folding reaction.

1.2 Protein Dynamics

1.2.1 Protein Folding

Protein folding is an intriguing biological problem with significant societal impacts. Protein misfolding and aggregation have been implicated in several diseases, including Alzheimer's, Parkinson's, and prion diseases [1]. A rigorous understanding of protein folding mechanisms would enable development of new pharmaceuticals to treat or prevent these diseases. For example, a drug developed by Genentech for prevention of Altzheimer's disease is currently in trial [2]. The drug is designed to target plaques of the protein beta amyloid and their oligomeric precursors that are thought to cause neurodegeneration.

Protein folding is also, in and of itself, an interesting biophysical problem. Proteins are heterogeneous polymers with unique primary amino acid sequences that dictate the secondary and tertiary structure, or the local and global intramolecular contacts, respectively. Many proteins fold *in vitro*, or outside of the cell,

from an ensemble of unfolded states to a single native state without the help of chaperones. Scientists are attempting to understand this process from multiple angles—through structure prediction from amino acid sequences [3, 4], folding kinetics and equilibrium measurements [5], computational modeling of protein dynamics [6], and unfolded state dynamics models and experiments.

The energy landscape theory [7, 8, 9, 10], developed by Wolynes, Bryngelson, and Onuchic, is a conceptual framework for the energetics of protein folding. In contrast to a simple free energy diagram with a discrete reaction pathway, multiple folding pathways are available for a protein to reach its global minimum, the native (folded) state. The principle of minimal frustration [11] describes how proteins with numerous accessible states often can fold efficiently to the native state, and why some proteins do not.

In minimally frustrated proteins, contacts that stabilize the native structure are strongly favored during folding, so that molecules appear to be funneled to the native state. Each newly formed native contact reduces the conformational search for the native structure. Frustration in the energy landscape will appear as kinetic traps, or energy barriers that can slow folding. A rough landscape, in which non-native contacts are energetically competitive with native contacts, can lead to deviations from simple first-order kinetics. Intermediates arise when there is a build-up of conformations due to a large energy barrier. If region(s) of the protein are more frustrated than others, a sequential folding mechanism with region-specific folding kinetics can result. It is, thus, necessary to probe multiple

regions of the protein for a global mechanism.

Although predicting folding mechanisms is very challenging, some important factors guiding folding have been identified [12, 13, 3]. These factors include topology (protein fold), the number of amino acids between tertiary contacts, chain length, and properties of amino acid sidechains, e.g., hydrophobicity. It is difficult to determine the impact of each parameter as they are interdependent. Therefore, we will compare proteins to other members of their family that share a common fold.

1.2.2 Intrachain Diffusion

We investigate tertiary contact formation rates in unfolded cytochrome cb_{562}. As formation of tertiary contacts is a necessary step for folding, the folding speed limit is set by the conformational dynamics in the unfolded state [14, 15, 16, 17]. When studying unfolded dynamics, frustration manifests as internal friction, slowing intrachain diffusion in regions of the polypeptide.

The Gray Group previously employed electron transfer quenching to estimate contact formation rates in denatured cytochrome *c* [15] and in α-synuclein [18, 19], the polypeptide implicated in Parkinson's disease. We have now extended our work on contact dynamics to include denatured cytochrome cb_{562}. Adiabatic electron transfer requires a close-contact encounter complex. If the electron transfer reaction is diffusion-limited, then the electron transfer rate constant corresponds to contact formation. Photosensitizers are used to probe the kinetics

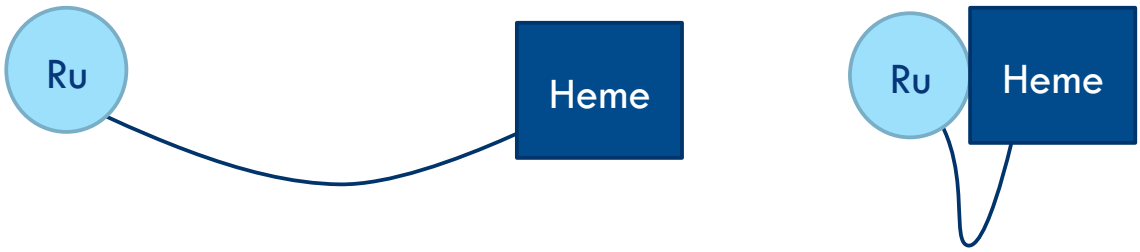

Figure 1.1: Contact quenching to probe intrachain diffusion. We examine formation of transient contacts between the heme and $[Ru(bpy)_2(IA\text{-}phen)]^{2+}$ complexes that have been covalently attached at various positions in unfolded cytochrome cb_{562}.

of intrachain diffusion (Figure 1.1).

1.2.3 Electron Tunneling

Electron transfer in metalloproteins is an important phenomenon in biology that occurs by quantum mechanical tunneling. Electron tunneling through protein medium occurs in nanoseconds to microseconds when there is a high driving force. We are interested in the factors that determine the electron transfer rate, especially the tunneling distance and the types of bonds in the pathway. We probe the kinetics of electron tunneling in folded cytochrome cb_{562} from a ruthenium photosensitizer, attached to various positions in the protein, to the heme using time-resolved spectroscopy (Figure 1.2).

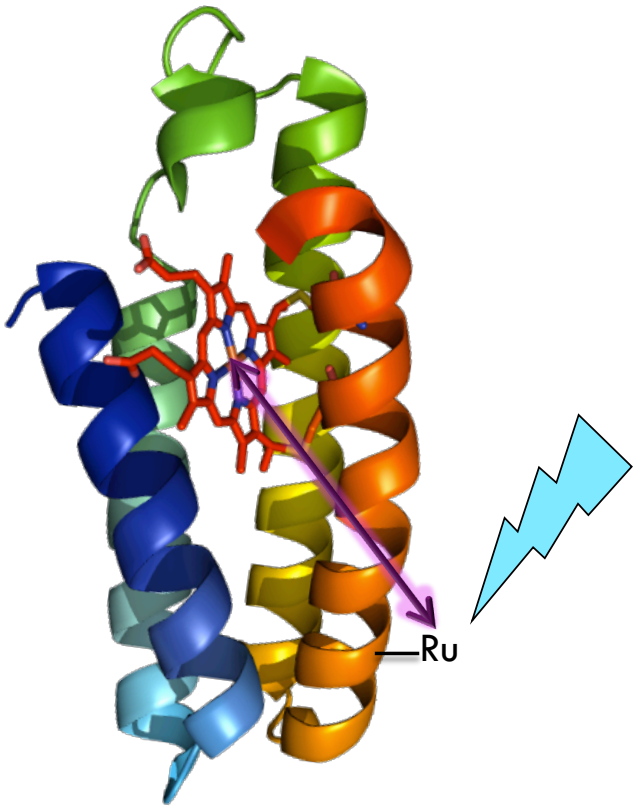

Figure 1.2: We measure ruthenium-heme electron tunneling rates in cytochrome cb_{562} following photoexcitation of $[Ru(bpy)_2(IA\text{-}phen)]^{2+}$ complexes that have been covalently attached at various positions in the protein.

1.3 Techniques to Probe Protein Dynamics

1.3.1 Time-Resolved FRET

Fluorescence energy transfer, or FRET, provides information about the distance, or the distribution of distances for an ensemble between two fluorophores. We use this technique to measure intramolecular distances in proteins and characterize heterogeneous ensembles of conformations, providing a "snapshot" of the different populations present at a point in time.

FRET is a nonradiative transfer of energy between two fluorophores that can occur when there is a spectral overlap between donor fluorescence and acceptor absorption. The efficiency of this process is described by Förster's model of dipole-dipole energy transfer: $E = {R_0}^6 / ({R_0}^6 + {r_{DA}}^6)$, where r_{DA} is the distance between the fluorophores and R_0 is the Förster distance, dependent on the spectral overlap and the dipole orientations [20].

In time-resolved FRET (trFRET), energy transfer is observed as a decrease in the lifetime of the donor fluorescence. The fluorescence decay is fit to an exponential function with each decay rate constant k corresponding to a donor-acceptor (D-A) distance (Figure 1.3). A single-exponential decay would suggest a homogeneous solution, whereas a multiexponential decay would indicate that there is a heterogeneous ensemble of molecules in different conformations.

Typically, proteins must be modified to incorporate fluorescent probes for FRET studies. (The fluorescent amino acid tryptophan can also serve as a probe.)

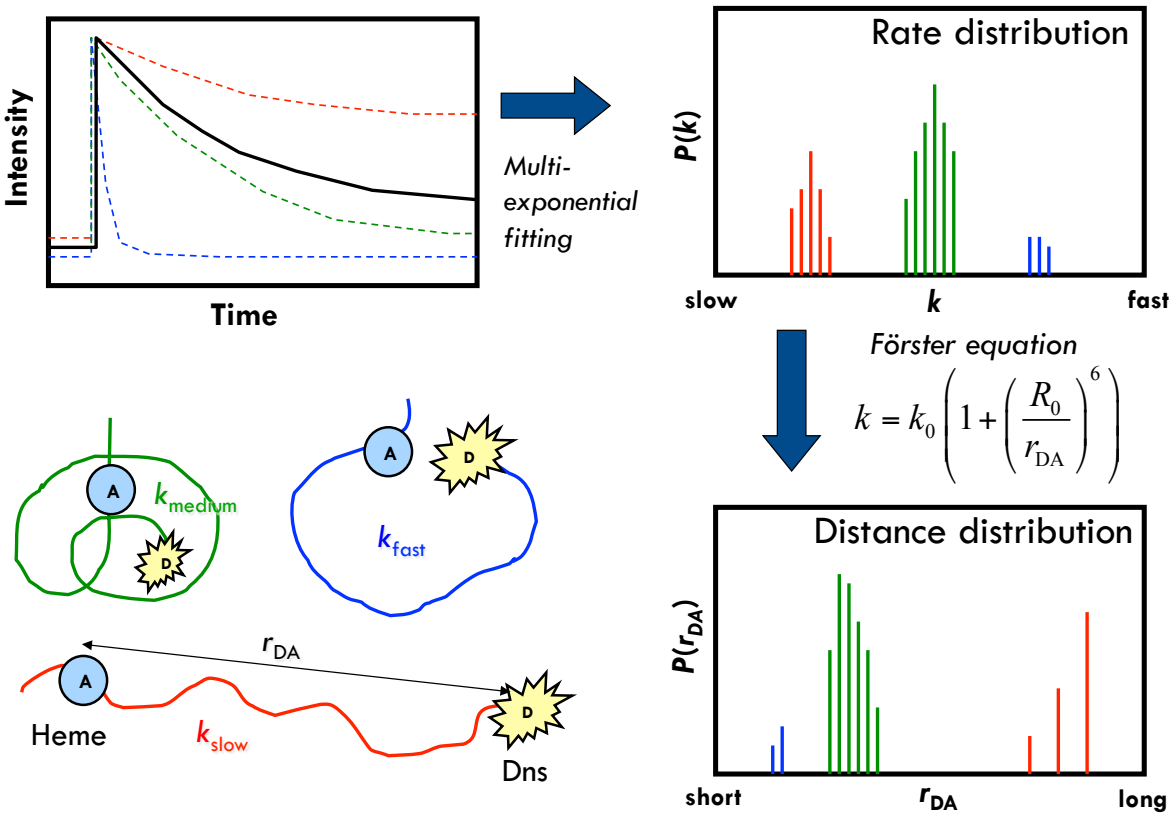

Figure 1.3: Time-resolved fluorescence energy transfer: Illustration of data analysis and results for a solution containing compact, intermediate, and extended protein conformations. The three populations contribute to the multiexponential luminescence decay, shown in black. Image courtesy of Seiji Yamada.

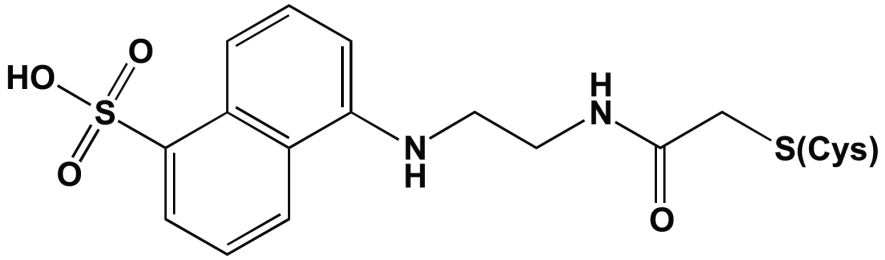

Figure 1.4: Dansyl fluorophore used in FRET studies.

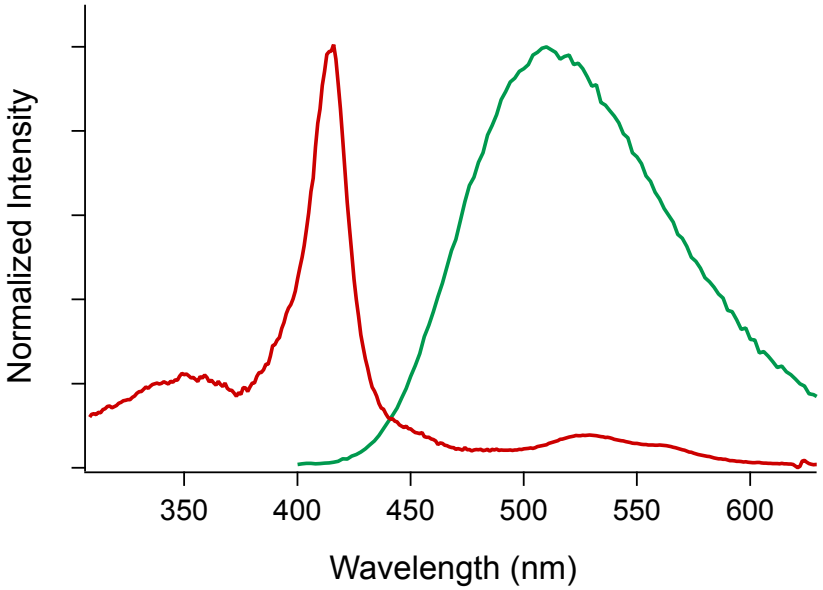

Figure 1.5: Spectral overlap of dansyl fluorescence (green, 355 nm excitation) and cytochrome cb_{562} heme absorbance (red).

We utilize the small organic fluorophore dansyl, attached covalently to cysteine residues, as the donor (Figure 1.4). The heme in the protein is the acceptor in our system. The dansyl-heme spectral overlap is shown in Figure 1.5. The Dns-heme pair has a Förster critical length (R_0) of 39 Å [21], which is optimum for measuring distances from 12–60 Å.

1.3.2 Continuous Flow Mixing to Trigger Folding

Protein folding is triggered by a change from conditions that disfavor folding, e.g., high temperature, acidic pH, or denaturants, to conditions that favor folding. Chemical denaturants, including guanidine hydrochloride (Gdn) and urea, are proposed to denature proteins by interacting with and disrupting the hydrogen bonding of the protein backbone and/or altering the intermolecular bonding of water [22, 23]. To trigger folding, denatured protein is mixed rapidly with buffer to decrease the denaturant concentration. In stopped flow mixing, the two solutions are mixed into a cell with high pressure. The "dead time," or time needed for complete sample mixing, of these instruments ($\sim$ 1–3 ms [24]) does not allow for resolution of fast folding processes. Newer models with smaller volumes and pneumatic drives reach dead times of 0.5–1 ms.

In contrast, continuous flow mixing allows for the resolution of submillisecond folding, including burst phases and early intermediates [25]. In continuous flow mixing, two solutions are rapidly and continuously mixed together in a microfluidic channel [26, 27]. The Gray Group has utilized continuous flow mixing to reveal the early phases of the folding of four-helix bundle cytochromes cb_{562} and c' [28, 29]. In this thesis, we resolve the fast formation of native cytochrome cb_{562} and the folding intermediate of cytochrome c_{552} by ultrafast continuous flow mixing.

We couple the mixer (150 μs dead time) with a picosecond streak camera to simultaneously monitor multiple mixer positions, corresponding to time points in

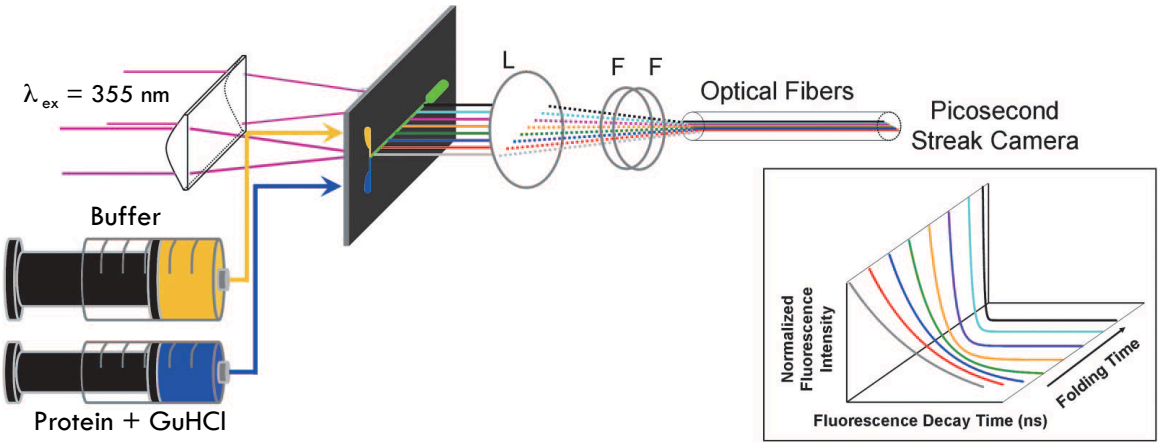

Figure 1.6: A continuous flow mixer is coupled to trFRET measurements. Folding is initiated upon mixing (green) of the unfolded protein (blue) and buffer (yellow). As the solution travels down the microfluidic channel, the protein has more time to fold; each position corresponds to a time point in the folding reaction. The excitation beam is focused on the channel, and fluorescence decays are collected with a picosecond streak camera. The detector allows for simultaneous data collection for multiple mixer positions (folding times). Modified with permission from Kimura et al. [28].

the folding reaction (Figure 1.6). The collected fluorescence decays (trFRET data) provide information about the changing distributions of protein conformations during folding, including the heterogeneity of structural ensembles. Measurement of a single donor-acceptor pair reveals the timescale for the formation of the structure in the corresponding regions of the protein. Fitting the time course of the mean donor-acceptor distance to an exponential function elucidates whether there is two-state folding with a single energetic barrier or if there are kinetic intermediates. Comparing the rate constants for two donor-acceptor pairs indicates whether the regions fold simultaneously or sequentially. Insights about folding mechanisms, whether there is a cooperative folding mechanism, a single sequential pathway, or multiple pathways, can be gleaned from this method.

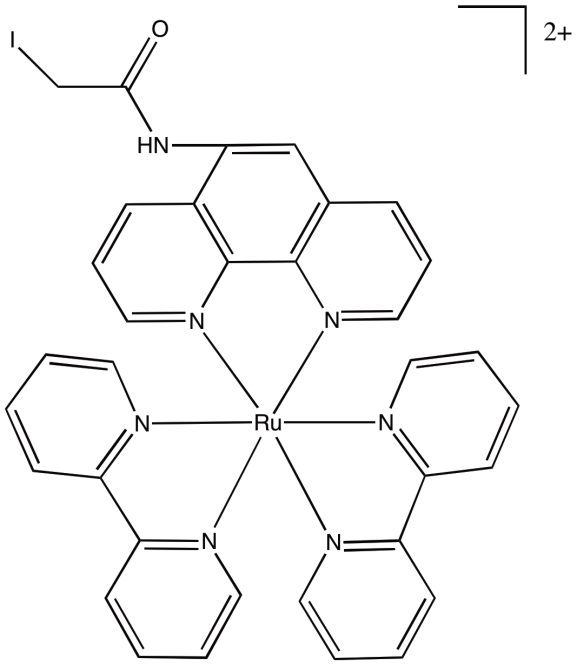

Figure 1.7: Ruthenium photosensitizer: $[Ru(bpy)_2(IA\text{-}phen)]^{2+}$.

1.3.3 Photochemical Triggers of Electron Transfer

Long-lived excited state metal complexes enable us to study electron transfer reactions on the microsecond timescale. The Gray Group has developed ruthenium complexes to probe electron transfer reactions in proteins [30]. We attach the complex $[Ru(bpy)_2(IA\text{-}phen)]^{2+}$ (Figure 1.7) to proteins site-specifically at mutated cysteines [31]. Electron transfer to the protein's heme occurs via adiabatic contact quenching or nonadiabatic electron tunneling. This reaction has a high driving force and is essentially activationless [32]. We photoexcite the Ru label in the MLCT absorption band (Figure 1.8) and observe electron transfer by monitoring the quenching of its luminescence or the absorption of transient species (Ru^{III} and Fe^{II}, Figure 1.9).

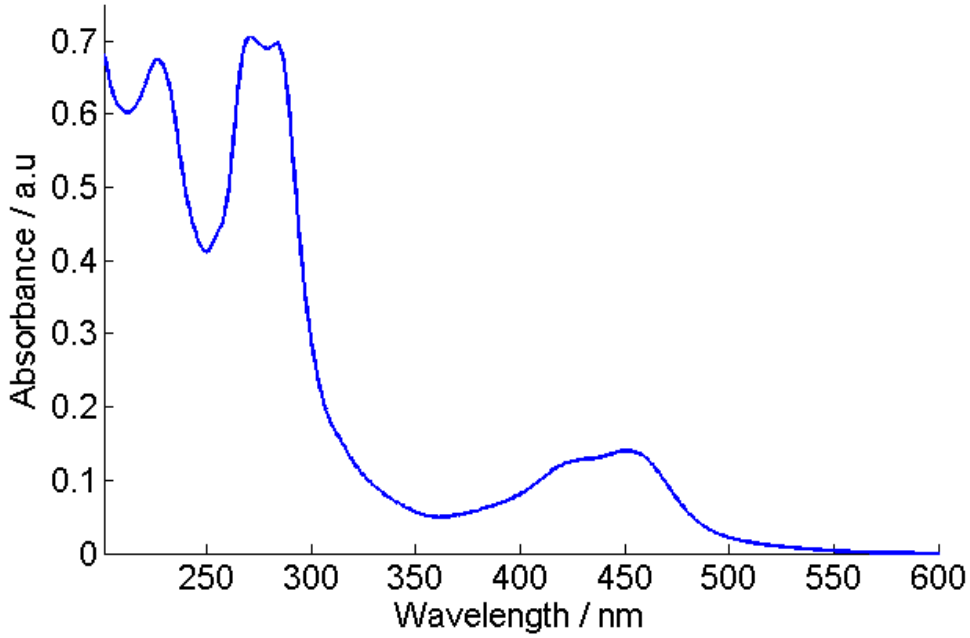

Figure 1.8: Absorption spectrum of $[Ru(bpy)_2(IA\text{-}phen)]^{2+}$. The label is photoexcited at 480 nm where there is minimal overlap with the heme Soret absorbance.

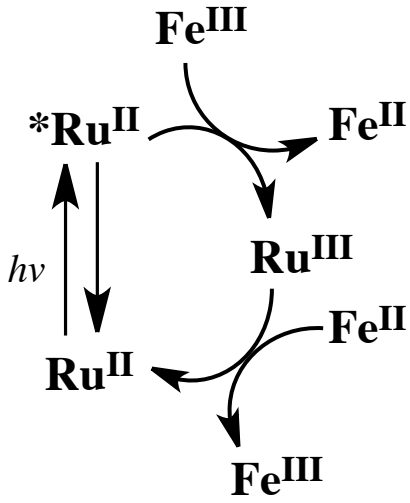

Figure 1.9: Ruthenium photochemistry: Triggering electron transfer to heme Fe^{III}.

1.4 Proteins of Interest

1.4.1 Cytochrome cb_{562}, a Four-Helix Bundle Cytochrome

The Gray Group studies the folding mechanisms of a family of four-helix bundle cytochromes, including cytochromes c', c_{556}, and b_{562} [33, 34, 35, 36, 37, 38, 28, 29, 39, 40]. While protein families typically have fairly conserved sequences and topologies, the members of this family are unique in that they have almost superimposable structures (Figure 1.10) [41, 42, 43] but unusually low sequence homology of only 10-20% [44]. Similar topology leads one to predict that the proteins should have similar folding kinetics; however these the four-helix bundle cytochromes have widely divergent folding kinetics. This rare characteristic creates a unique opportunity to study the relationship between amino acid sequences and folding pathways in proteins with similar topologies.

The folding kinetics of cytochrome c' are highly diverse, with time constants ranging from milliseconds to seconds [28, 38]. Probe-dependent folding kinetics and heme transient absorption spectra reveal deviations from two-state folding behavior and a sequential folding mechanism. Ferricytochrome c' has heterogeneous conformations during refolding, including a population that is more compact than the native state, suggestive of hydrophobic collapse during folding [28]. The large degree of frustration is partially attributed to proline isomerizations [38]. Folding of ferrocytochrome c' is further slowed by methionine misligations of the heme [34].

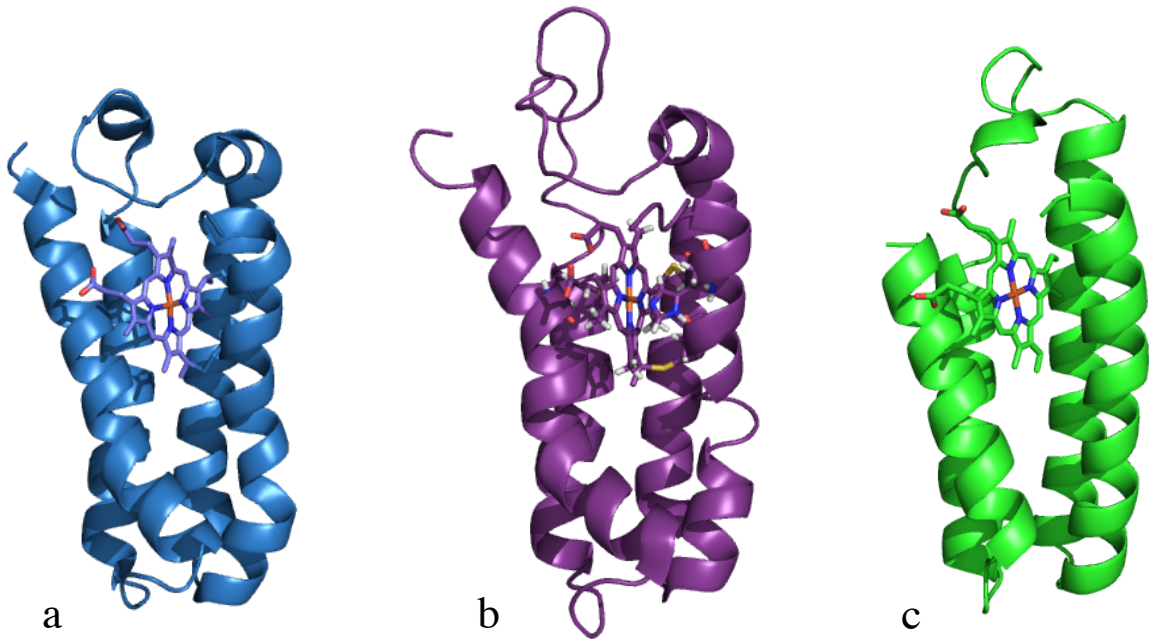

Figure 1.10: Stuctures of four-helix bundle cytochromes. Structures cytochrome c′ and cytochrome b_{562} have a backbone atom RMSD of 3.4 Å, but only 15% sequence homology. Despite their conserved topology, the four-helix bundle cytochromes have highly divergent folding kinetics. (a) Cytochrome *c′* from *R. palustris* (PDB 1A7V). (b) Cytochrome c_{556} from *R. palustris* (PDB 1S05). (c) Cytochrome b_{562} from *E. coli* (PDB 256B).

On the other hand, cytochrome b_{562} folds the fastest among the four-helix bundle cytochromes with simple first-order folding kinetics [33]. Folding of ferrocytochrome b_{562} was triggered by electron injection to the oxidized iron center in slightly denaturing conditions, taking advantage of stability differences between the ferric and ferrous oxidation states. Folded protein was formed on a submillisecond timescale; however, folding rate determination was limited by irreversible heme dissociation. Furthermore, evidence suggests that the measured rate constant may only represent the folding event for a small population of a native-like conformation present in low concentrations of denaturant [33, 45].

To overcome the complication of heme dissociation during refolding, cytochrome b_{562} with its heme bound solely by heme-ligand interactions was modified to cytochrome cb_{562} with additional *c*-type covalent heme linkages to the protein backbone [37]. Cytochromes *c′* and cb_{562} now share a common topology and similar heme environment so that the influence of the primary sequences on folding kinetics can be isolated from these factors. Cytochrome cb_{562} has been shown by Trp59-heme trFRET studies to fold on the millisecond timescale with simple first-order folding kinetics [29]; we further characterize cytochrome cb_{562}'s folding mechanism in this thesis.

Cytochrome cb_{562} is an excellent model α-helical protein. It is very stable and refolds *in vitro*. Due to its high degree of symmetry, it has been utilized as a building block for macromolecular protein assemblies [46]. In addition, it is thought to be an electron transfer protein, although its specific function is

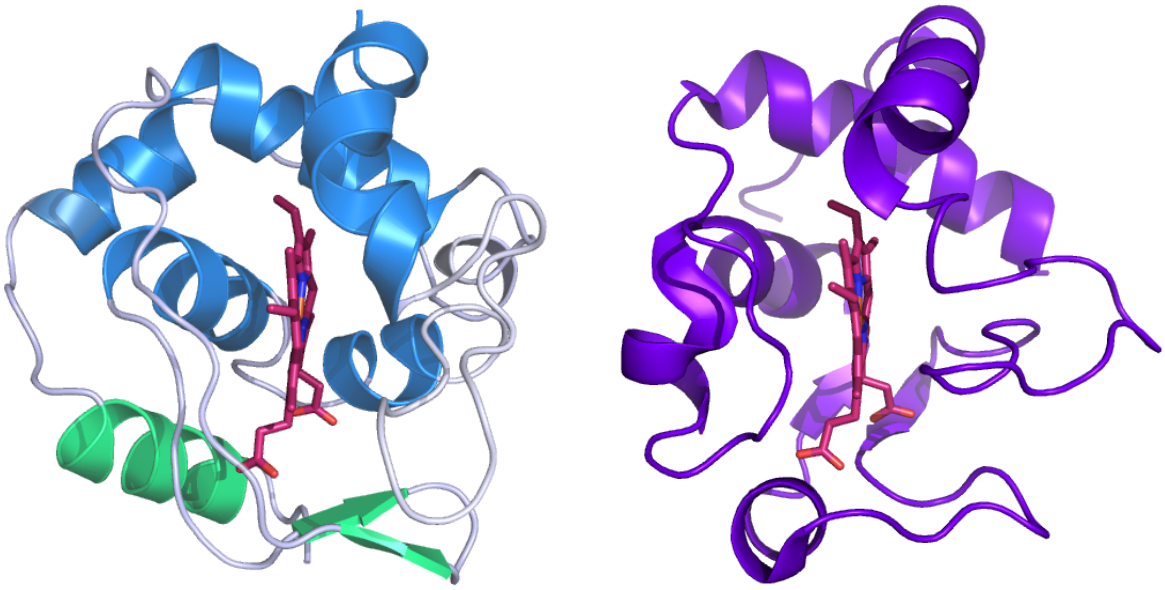

Figure 1.11: Structures of class-I c-type cytochromes c_{552} from *T. thermophilus* (blue/green, PDB 3VNW) and yeast *c* (violet, PDB 1YCC). The N- and C-terminal helices cross each other (apexes). The structural features in green are distinct to cytochrome c_{552}.

unknown. We have shown that Zn-cytochrome b_{562} is photoconductive [47], and in this thesis, we report electron tunneling rates for Ru-modified cytochrome cb_{562}.

1.4.2 Cytochrome c_{552}, a Class I C-Type Cytochrome

Cytochrome c_{552} is a small soluble, thermostable protein containing a single heme bound covalently via two cysteines and axially ligated by His15 and Met69. Cytochrome c_{552} is found in the hyperthermophile *Thermus thermophilus* with an optimal growth temperature of $\sim$65°C that was discovered in a Japanese hot spring. It is the natural electron donor for cytochrome ba_3 oxidase during the terminal step of respiration [48].

Cytochrome c_{552} belongs to the class-I *c*-type cytochrome family, which includes

yeast cytochrome *c*. Cytochromes in this family share the common structural feature where the N- and C-terminal helices cross each other (Figure 1.11). However, cytochrome c_{552} has the distinct structural components of a β-sheet around the heme propionate and extra helices in the C-terminal region (Figure 1.11, green region). We are interested in the role of these distinct structural features and their possible correlation to the thermostability of the protein.

Brunori and coworkers examined the folding pathway of this protein by native tryptophan (Trp91) fluorescence kinetics and revealed that there is an on-pathway folding intermediate [49]. In this work, Trp91 probed structures near the crossed helices of cytochrome c_{552}, but did not report on the heme β-sheet or the C-terminal helices. We would like to further characterize the folding intermediate structurally.

1.5 Thesis Overview

We gain insight into cytochrome cb_{562}'s dynamics and folding mechanism in Chapters 3–5. In Chapter 3, we spectroscopically characterize cytochrome cb_{562}'s structural conformations, heme environment, and stability to chemical denaturants. In Chapter 4, we probe the unfolded dynamics and intrachain diffusion of cytochrome cb_{562} by electron transfer. In Chapter 5, we investigate the folding kinetics of cytochrome cb_{562} by trFRET and microfluidic mixing.

We explore the folding of cytochrome c_{552} in Chapters 6–7. We characterize an equilibrium unfolding intermediate of cytochrome c_{552} by trFRET and improve our understanding of the physical basis for its thermostability in Chapter 6. In Chapter 7, we resolve cytochrome c_{552}'s folding intermediate by trFRET and microfluidic mixing.

The findings from the protein folding studies are summarized in Chapter 8. Finally, we investigate the effects of structural changes on electron tunneling rates in cytochrome cb_{562} (Chapter 9).

Chapter 2

Methods

2.1 General

Buffers were prepared with MilliQ water and filtered (0.22 μm). Chemicals were purchased from Sigma unless stated otherwise. Guanidine hydrochloride (Gdn, Sigma Ultra) concentration was determined by refractive index measurements [50].

2.2 Protein Preparation

2.2.1 Cytochrome cb_{562} Expression and Purification

Mutations were introduced in the cyt cb_{562} plasmid [37] by site-directed mutagenesis using a QuikChange kit (Stratagene) with oligonucleotide primers purchased from Eurofins MWG Operon; e.g., the K83C forward primer was GCTGGCA AATGAAGGT**TGC**GTAAAAGAAGCGC. After the polymerase chain reaction, reactions were digested with the restriction enzyme DpnI to remove

parental methylated DNA. The mutant DNA was amplified by transformation into *E. coli* XL1-blue cells (Stratagene) and expressed overnight at 37°C on LB agar plates containing 100 μg/mL ampicillin. LB medium with 100 μg/mL ampicillin or carbenicillin was inoculated with colonies, and the cells were expressed at 37°C with shaking for 14-16 h. Plasmids were then isolated using a Qiagen Miniprep kit, and the mutations were confirmed by DNA sequencing (Laragen, Inc.).

The cyt cb_{562} plasmid was cotransformed with a cytochrome *c* maturation gene cassette pEC86 in BL21 Star (DE3) One Shot *E. coli* (Invitrogen), as previously described [37]. Transformants were grown overnight at 37°C on LB agar plates containing 100 μg/mL ampicillin and 34 μg/mL chloramiphenicol. Single colonies were selected to inoculate LB medium with the same concentration of antibiotics and shaken for 6-8 h. Glycerol stocks of the cultures were prepared by flash freezing a 3:1 mixture of culture to 80% glycerol using liquid nitrogen and stored at -80°C. The remaining test cultures were incubated for an additional 10-12 h, and then the cells were harvested by centrifugation. Bright red colonies were indicative of cyt cb_{562} overexpression; the corresponding glycerol stocks were selected for subsequent protein expressions.

Proteins were expressed by *E. coli* in LB medium containing 100 μg/mL ampicillin and 34 μg/mL chloramphenicol. Small cultures (3035 mL) were inoculated with glycerol stocks, shaken for 8 hours at 37°C, and distributed into large cultures (12 L). After expressing for 16 h at 37°C with rotary shaking, cells were harvested and stored at -80°C. Protein was extracted from cells by sonicating

for 5 s pulses at 50% duty cycle for 30 min in Tris-HCl buffer (pH 8.0) containing 2 mM ethylenediaminetetraacetic acid, 13 mM dithiothreitol (DTT), 1 mg deoxyribonuclease I, 6 mM sodium deoxycholate, and 7 mM phenylmethanesulfonyl fluoride, predissolved in isopropyl alcohol with dimethyl sulfoxide. Yield was higher when cells were lysed within a few days of expression.

Following centrifugation, the supernatant containing the crude protein was purified on a Q-Sepharose Fast Flow column (GE Healthcare) in 10 mM Tris-HCl (pH 8.0) with 1 mM DTT and a typical gradient of 0 to 110-150 mM NaCl. Cyt cb_{562} was subsequently purified by fast-protein liquid chromatography on a Pharmacia AKTA Purifier system. Disulfide bonds were reduced with 5-20 mM DTT, and the buffer was exchanged with 15 mM sodium acetate buffer (pH 4.5) on a HiPrep 26/10 desalting column (GE Healthcare). Purified protein was eluted with a NaCl gradient on a Mono-S 10/10 GL cation-exchange column (GE Healthcare). Depending on the mutant, the protein was eluted with a final concentration of 130 to 150 mM NaCl. When necessary, the protein was further purified on a Mono-Q 10/10 GL anion-exchange column (GE Healthcare) at pH 8.0 in 10 mM Tris-hydrochloride buffer (pH 7) with a NaCl gradient. Samples were concentrated using an Amicon YM10 membrane (Millipore).

2.2.2 Cytochrome c_{552} Expression and Purification

Cyt cyt c_{552} variants were prepared by Seiji Yamada, a visiting scientist from Sony Corporation. For detailed methods, Supporting Information is free on the

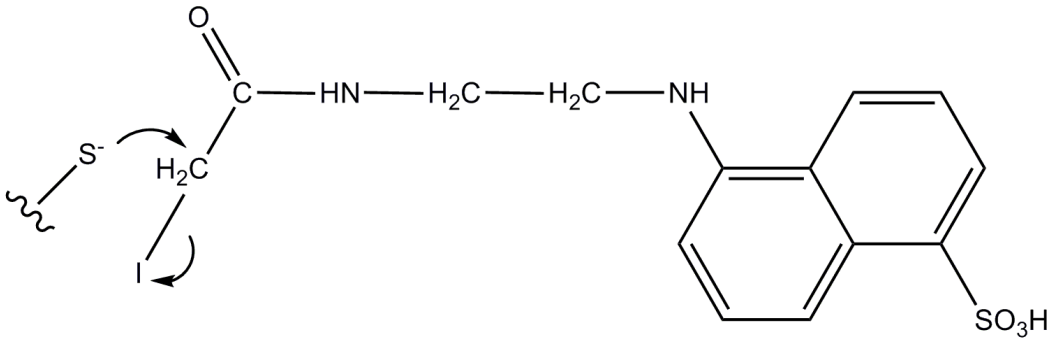

Figure 2.1: Covalent labeling of cysteine residues with dansyl.

web: www.pnas.org/ content/110/5/1606/suppl/DCSupplemental [51].

2.2.3 Covalent Labeling with Photosensitizers

Photosensitizers were covalently attached to cysteine thiols in labeling reactions (Figure 2.1). For trFRET measurements, single-cysteine variants were labeled with a dansyl group (1,5-IAEDANS, 5-((((2iodoacetyl)amino)ethyl)amino) naphthalene-1-sulfonic acid, Invitrogen). For electron transfer luminescence measurements, single-cysteine variants were labeled with $[Ru(bpy)_2(IA\text{-}phen)]^{2+}$ (IA-phen = 5-iodoacetamido-1,10-phenanthroline), synthesized by members of the Gray Group as described previously [31].

Disulfides bonds were reduced with 20 mM dithiothreitol (DTT) at pH 8; concentrated Tris-hydrochloride buffer used to adjust the pH as needed. The protein was exchanged into pH 8.0 buffer on a HiPrep 26/10 desalting column (GE Healthcare) or a PD-10 desalting column (GE Healthcare). Tris-hydrochloride buffer was used for most labeling reactions; however, many dansyl-labeling reactions were carried out in 6 M Gdn with 50 mM HEPES (pH 8.0). Molar excess

of the label (5–10x dansyl label, 2–5x Ru label) was dissolved in dimethyl sulfoxide and slowly added to the Ar-deoxygenated protein. The sample was stirred in the dark for 3–4 h at room temperature or overnight at 4°C.

The reaction was quenched with DTT, and the buffer was exchanged with 15 mM sodium acetate buffer (pH 4.5) using the HiPrep 26/10 desalting column. Labeled protein was separated from unlabeled protein on a Mono-S 10/10 GL column (GE Healthcare) equilibrated with 15 mM sodium acetate buffer (pH 4.5) or a Mono-Q 10/10 GL column (GE Healthcare) equilibrated with 10 mM Tris-hydrochloride buffer (pH 8.0) using a NaCl gradient. Samples were concentrated using an Amicon Ultra 3 kDa unit (Millipore). As the photosensitizers are light-sensitive, all labeled samples were kept in the dark.

2.3 Protein Characterization

Purity of the expressed and labeled protein variants was verified by SDS-PAGE on a Pharmacia PhastSystem (Amersham Biosciences). The molecular mass of the expressed protein was confirmed by electrospray ionization mass spectroscopy (Caltech Protein/Peptide Microanalytical Laboratory). Circular dichroism (CD) and fluorescence spectra were collected using an Aviv 62ADS spectropolarimeter (Aviv Associates, Lakewood, NJ) and a Fluorolog2 spectrofluorimeter (Jobin Yvon), respectively.

The purity, concentration, and heme environment of cyt cb_{562} and c_{552} samples were assessed by UV-visible spectroscopy using an Agilent 8453 diode array spec-

trophotometer (Agilent Technologies, Santa Clara, CA). The extinction coefficient for oxidized cyt cb_{562} is 0.148 $\mu M^{-1}cm^{-1}$ at the maximum Soret absorbance (415 nm) [37]; to confirm the purity of each sample, an absorbance ratio (418 nm: 280 nm) of five or greater was verified. The extinction coefficient for reduced cyt c_{552} is 21.1 $mM^{-1}cm^{-1}$ [52].

Though mutant positions were carefully selected at solvent-exposed positions, stability studies were performed to quantify perturbation by the cysteine mutation and the addition of a dansyl group. Denaturant unfolding curves of the proteins were obtained by fitting CD, absorbance, and fluorescence spectroscopy data to a two-state unfolding mechanism [22] using Igor Pro version 5.02 or 6.1 (Wavemetrics). Samples of 2 μM cyt cb_{562} were prepared in 50 mM sodium acetate (pH 5.0) with Gdn (0–6 M); the protein was allowed to equilibrate in denaturing conditions for an hour. The unfolding curves were generated from CD (222 nm), absorption (402 nm: 415 nm), and Trp59 fluorescence (355 nm excitation and integration of emission) data. Samples of 2-20 μM cyt c_{552} were measured in various concentrations of Gdn buffered with 100 mM sodium citrate (pH 3.0) for absorbance and fluorescence measurements and with 10 mM citrate buffer for CD measurements; the unfolding curves were generated from CD (222 nm), absorption (391 nm), and Dns fluorescence (355 and 513 nm for excitation and emission, respectively) data.

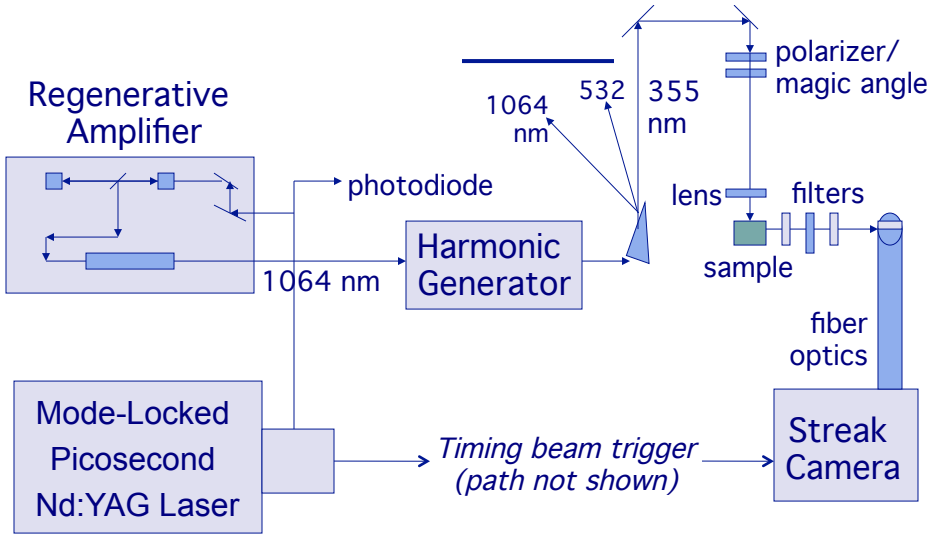

Figure 2.2: Schematic of the picosecond laser system.

2.4 Picosecond Laser System

For trFET measurements, samples with 1-2 μM protein were prepared with 0–8 M Gdn in buffer (cyt cb_{562}: 50 mM sodium acetate (pH 4–5) or 100 mM sodium phosphate buffer (pH 7); cyt c_{552}: 100 mM sodium citrate (pH 3.0)). Samples were deoxygenated by 30 pump-back-fill cycles with argon. Samples in denaturing conditions were allowed to equilibrate for at least an hour before data collection.

A schematic of the laser system is shown in Figure 2.2. Samples were excited at 355 nm with the third harmonic (10 ps, 0.5 mJ) from a regeneratively amplified (Continuum), mode-locked picosecond Nd:YAG laser (Vanguard 2000-HM532; Spectra-Physics). A long-pass cutoff filter (>430 nm) selected for dansyl fluorescence. Luminescence was collected using a picosecond streak camera (C5680; Hamamatsu Photonics) in the photon-counting and analog integration modes for equilibrium unfolding and refolding kinetics experiments, respectively.

The decay data were collected on short (5 ns) and long (20 or 50 ns) time scales. Single-shot streak camera traces (10,000 exposures for 5 ns and 5,000 exposures for 50 ns timescales) were averaged. Samples were stirred continuously. Data were collected at ambient temperature ($\sim$18°C).

The resulting short- and long-time scale data were spliced together; the point of maximum laser pulse intensity was defined as t=0. The combined traces were compressed logarithmically before fitting (70 points per decade); this compression does not alter the interpretation of data [29]. Fluorescence energy transfer analysis involves the numerical inversion of a Laplace transform [$I(t) = \sum_k P(k)\exp(-kt)$] [53] [54]. The trFET data were fit using MATLAB (Mathworks) with a Tikhonov regularization (TR) and a maximum-entropy (ME) fitting algorithm with regularization methods that impose additional constraints on the properties of $P(k) \geq 0$.

TR fitting minimizes the sum of the squares of the deviations between calculated and observed fluorescence intensities (χ^2), as well as the gradient of the $P(k)$ distribution ($\bigtriangledown P(k)$). ME fitting minimizes the sum of the squares of the deviations between calculated and observed fluorescence intensities (χ^2), while maximizing the entropy of the rate-constant distribution [$S = -\sum_k P(k)\ln[P(k)]$]. TR and ME fitting yield stable and reproducible numerical inversions of the kinetics data. The balance between minimization of χ^2 and $\bigtriangledown P(k)$ minimization or S maximization was determined by L-curve analysis [55].

The probability distributions of rate constant, $P(k)$, were converted to prob-

ability distributions of distance, P(r_{DA}), using the Förster equation [k = k_0 (1 + $(R_0/r_{DA})^6$)] [56]. The decay rate constant of the unquenched Dns fluorophore, k_0, is 1.0 x 10^8 s^{-1}, determined using Dns-modified cysteines. The Förster critical length, R_0, for the Dns-heme pair is 39 Å under both native and denatured conditions [21]. At distances longer than $1.5r_0$, energy transfer quenching of D is not competitive with excited-state decay; D-A distances cannot be obtained reliably, and different structures in the protein ensemble with $r \geq 59$ Å cannot be resolved.

2.5 Continuous Flow Mixer

For refolding kinetics experiments, Gdn-denatured protein was mixed with buffer (volume ratio of 1:5) in a T-shaped continuous flow mixer to initiate folding (Figure 1.6) [57]. In Chapter 5, denatured cyt cb_{562} in 6 M Gdn, 50 mM sodium phosphate (pH 4.0) was mixed with 50 mM sodium phosphate (pH 4.0). In Chapter 7, denatured cyt c_{552} in 6 M Gdn, 100 mM sodium citrate (pH 3.0) was mixed with 100 mM sodium citrate (pH 3.0). Final protein concentrations were 12-15 μM.

The mixing cell, based on a design by Fujisawa and coworkers [58], was given to the Gray Group by former group member Dr. Tetsunari Kimura. The device has a mixing dead time of $\sim$150 μs [57]. A 200-μm-thick stainless steel plate (Toray Precision, Shiga, Japan) was etched with two flow channels (35 μm width, 100 μm depth) and an observation channel (200 μm width and depth). The plate was placed between two synthetic fused silica windows and into a stainless steel holder. The layers were held in place by screw-tightened Teflon pieces.

The mixer was carefully assembled to avoid introduction of dust particles. Pieces were sonicated, rinsed with water, and dried with Kimwipes or pressurized air. The quartz can be further cleaned by soaking in dilute hydrochloric acid. Following assembly, the device was tested for leaking. Uninterrupted flow through the mixer was confirmed with a microscope. After use, the mixer and tubing were dried by injecting air. An online filter prevented particles from entering the mixer.

Two solutions were injected into the mixer by a syringe pump at a combined flow rate of 3 mL/min. The picosecond excitation beam (see section 2.4) was focused into a strip along the observation channel using a combination of spherical and cylindrical lenses. Luminescence from the mixing chamber was focused onto a fiber-optic bundle with optical fibers in a linear array, allowing for simultaneous data collection at twenty-five different points along the mixing chamber. At this flow rate (3 mL/min = 125 cm/s in mixer), the folding time resolution was 200 μs, calculated using a fiber-fiber distance of 250 μm. The streak camera was operated in the analog integration mode, and the background signal of buffer with the corresponding concentration of Gdn was collected and subtracted prior to analysis.

We performed control experiments by flowing protein in 1 M or 6 M Gdn through the mixer. We confirmed that the fluorescence decay for each fiber was consistent with that of the folded and unfolded proteins, respectively. Twenty fibers in the center provided the best signal.

The moments (M_n) of the P(r) distributions were calculated according to the equation [$M_n = \sum \mathrm{P}(r)r^n \,/\, \sum \mathrm{P}(r)$]. The variances ($V$), which reflect the breadth of the P(r) distributions, were calculated according to the equation [$V = M_2 - (M_1)^2$].

2.6 Nanosecond Laser System

For contact quenching studies (Chapter 4), luminescence decays were collected for Ru-labeled cyt cb_{562} (1.5–6 μM) denatured with 6 or 8 M Gdn in 50 mM sodium acetate buffer (pH 4–5) or 100 mM sodium phosphate buffer (pH 7). For bimolecular measurements, $Ru(bpy)_3Cl_2$ (1 μM) was combined with cyt cb_{562} (0–175 μM) in 6 M Gdn, 50 mM sodium acetate buffer (pH 4). For electron tunneling studies (Chapter 9), luminescence and transient absorption data were collected for Ru-labeled cyt cb_{562} (6 μM) in 100 mM sodium phosphate buffer (pH 7). Samples were deoxygenated by 30 pump-back-fill cycles with argon.

Luminescence and transient absorption data were collected with the setup illustrated in Figure 2.3. Samples were excited at 480 nm with 10 ns laser pulses from an optical parametric oscillator (Spectra-Physics Quanta-Ray MOPO-700), pumped by the third harmonic of a Q-switched Nd: YAG laser (Spectra-Physics Quanta-Ray PRO-Series) at a rate of 10 Hz. Transmitted light (luminescence or probe light) was detected by a photomultiplier, amplified, and digitized. The laser system was controlled by software written in LabVIEW (National Instruments).

For luminescence measurements, a long-pass cutoff filter (>600 nm) and a monochromator with 1 mm slits selected for Ru luminescence (630 nm). Lumi-

nescence decays comprised of three hundred averaged laser shots were collected on 2 μs or 10 μs timescales, as appropriate. Samples were stirred continuously.

For transient absorption measurements, probe light (75-W arc lamp, PTI model A 1010) was aligned with the excitation beam at the sample. Appropriate short-pass and long-pass filters and the monochromator selected for the wavelength. For measurements collected with the fast amplifier (<msec), the probe light was pulsed. Samples were stirred continuously, and six hundred traces were typically averaged. For measurements collected with the slow amplifier, samples were stirred between 20–40 cycles of one shot each.

Data were fit to exponential functions using the Curve Fitting Tool (cftool) in MATLAB version R2010b (Mathworks). The point of maximum laser pulse intensity was defined as t=0. Data in the first 10-15 ns that overlapped with the laser pulse were excluded from fits. At high protein concentrations, such as in the bimolecular experiments, a high intensity spike was observed in the luminescence at time zero with deviations in the signal for up to 30 ns; these data points were excluded from the fits. A fitting range of 3e3 to 3e8 was set for the rate constant.

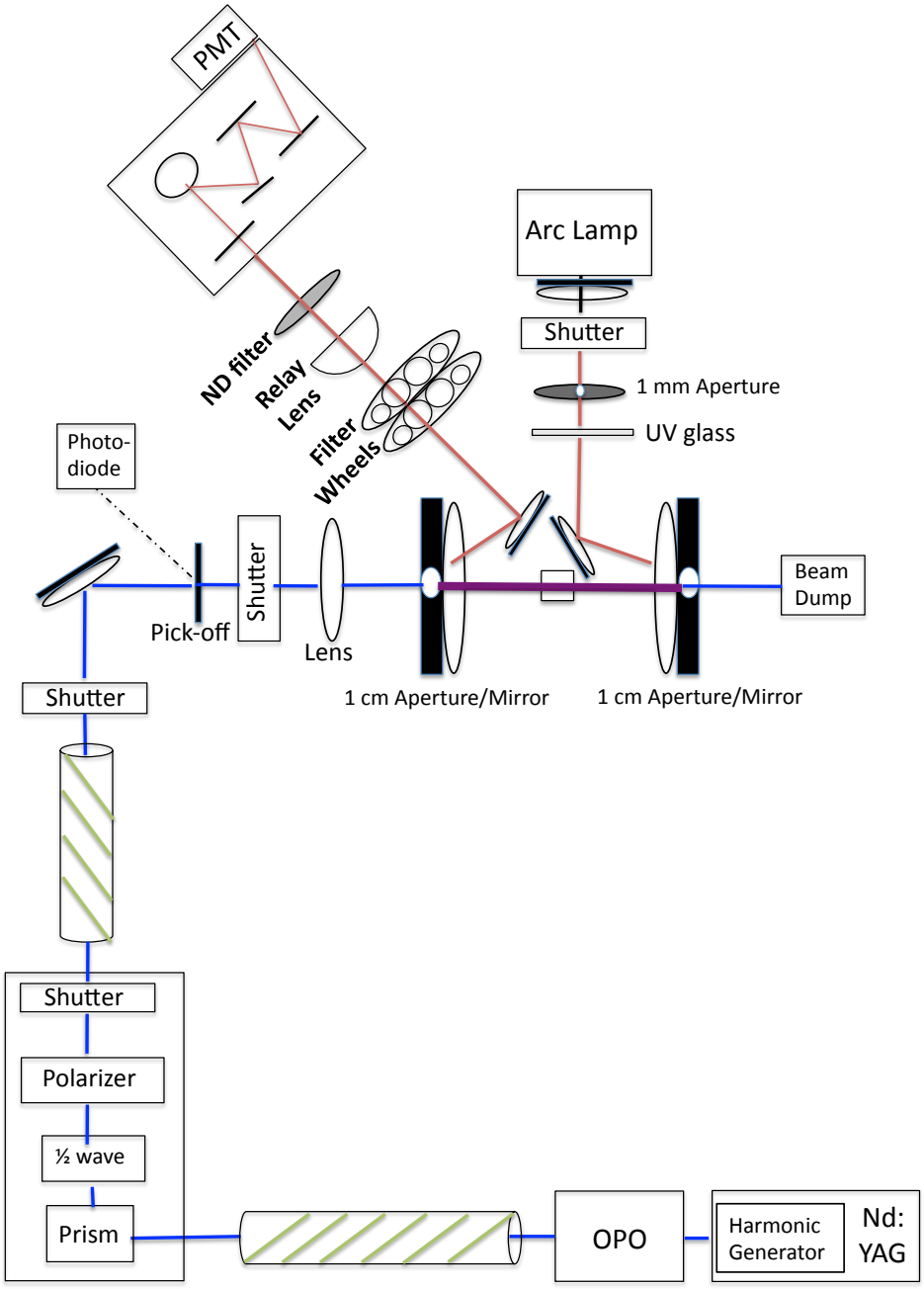

Figure 2.3: Schematic of the nanosecond laser system. The laser beam (blue) excites the sample. During transient absorption measurements, the difference in the absorbance of the probe light (red) is measured over time. While measuring the luminescence of the sample, the shutter for the probe light is closed. The wavelength is selected for using filters and a monochromator. Image courtesy of Gretchen Keller.

Chapter 3

Biophysical Characterization of Cytochrome cb_{562}

3.1 Introduction

To prevent heme dissociation during refolding, the porphyrin vinyl groups of cytochrome cb_{562}'s heme were covalently attached to two mutated cysteine residues, R98C and Y101C, through the conserved CXXCH cytochrome *c*-type motif [37]. The heme environment is not significantly perturbed from cytochrome b_{562}, and the average RMSD for α-carbons is 0.42 Å [41, 37]. An additional mutation, K59W, aids in expression and can be used as a fluorescent probe. Cytochrome cb_{562} folds reversibly and has enhanced stability in denaturant unfolding conditions (folding free energy change of -42 kJ/mol at 0 M Gdn) when compared to cytochrome b_{562} (-30 kJ/mol) [37].

In this chapter, we characterize native and unfolded cytochrome cb_{562}. We monitor how the heme environment changes with denaturation and pH changes by UV-visible absorption spectroscopy. We measure the stability of the protein to denaturation, and we then investigate the unfolded state ensemble conformations

by trFRET. Residues have been selected throughout the protein for labeling with a dansyl (Dns) fluorophore in single-cysteine variants.

3.2 Protein Purification and Characterization

Cytochrome cb_{562} was purified and the labeled protein was separated by ion-exchange chromatography on an FPLC system. A representative chromatogram is shown in Figure 3.1. At pH 4.5, Dns-labeled protein is less cationic than unlabeled protein, whereas Ru-labeled protein is significantly more cationic than unlabeled protein. The elution conductivity varies with mutations.

The folded structure and purity were confirmed by UV-visible adsorption spectroscopy, and the mass was verified by electrospray ionization mass spectrometry (ESI-MS). Representative mass spectra for the Dns-labeled (Chapters 3 and 5) and Ru-labeled (Chapters 4 and 9) proteins are shown in Figures 3.2 and 3.3.

3.3 Stability of Labeled Mutants to Denaturation

Solvent-exposed residues that form no intramolecular bonds were selected for photosensitizer labeling to minimize structural perturbations. To confirm that the mutations and Dns labels do not destabilize the protein, denaturation curves were obtained by circular dichroism (CD) and UV-visible spectroscopy. The relative populations of folded and unfolded protein when the protein is denatured at various concentrations of guanidine hydrochloride (Gdn) were measured.

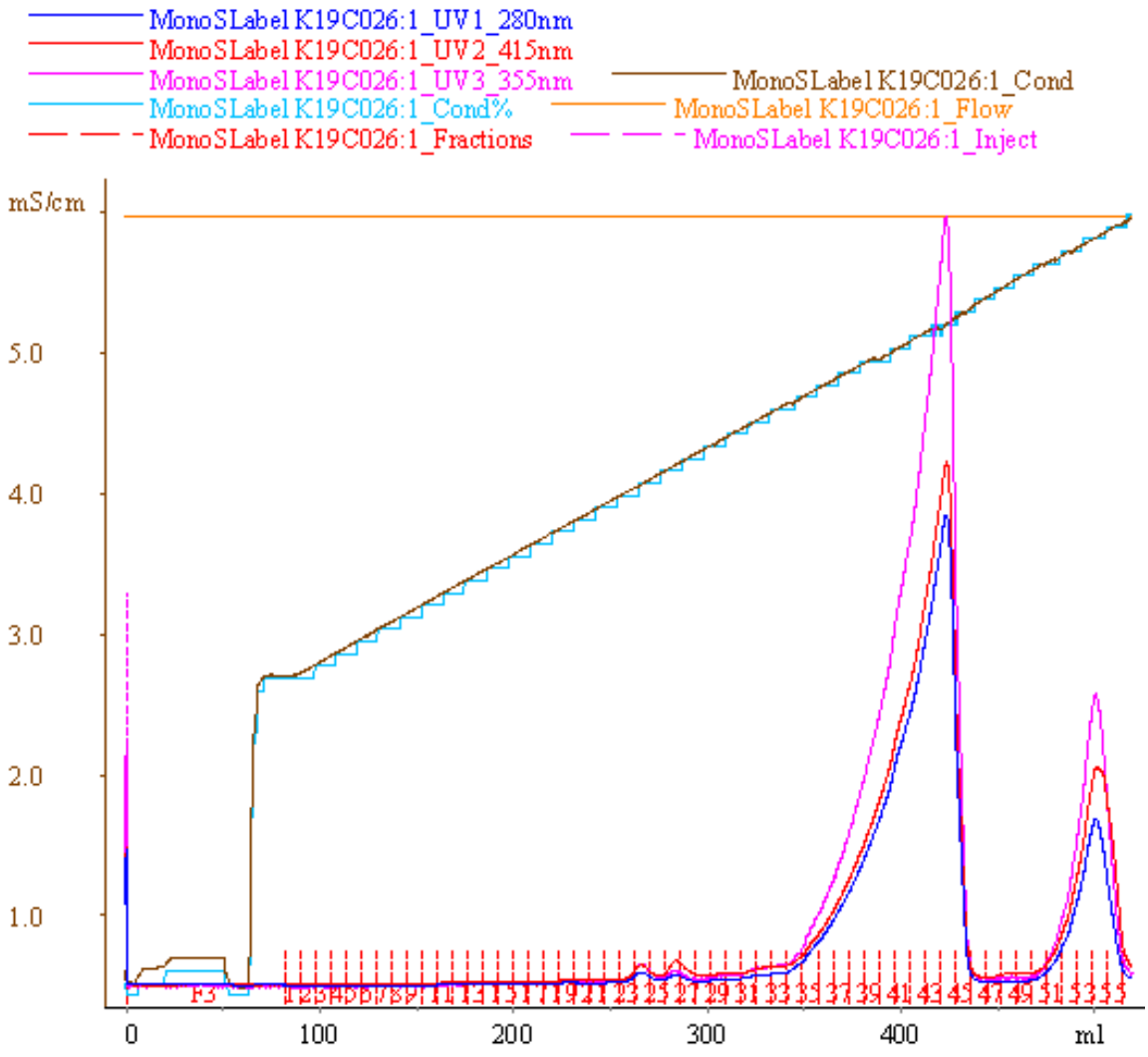

Figure 3.1: Representative FPLC chromatogram. Separation of Dns-labeled (first peak) and unlabeled (second peak) cytochrome cb_{562} (variant K19C) by cation-exchange chromatography (pH 4.5).

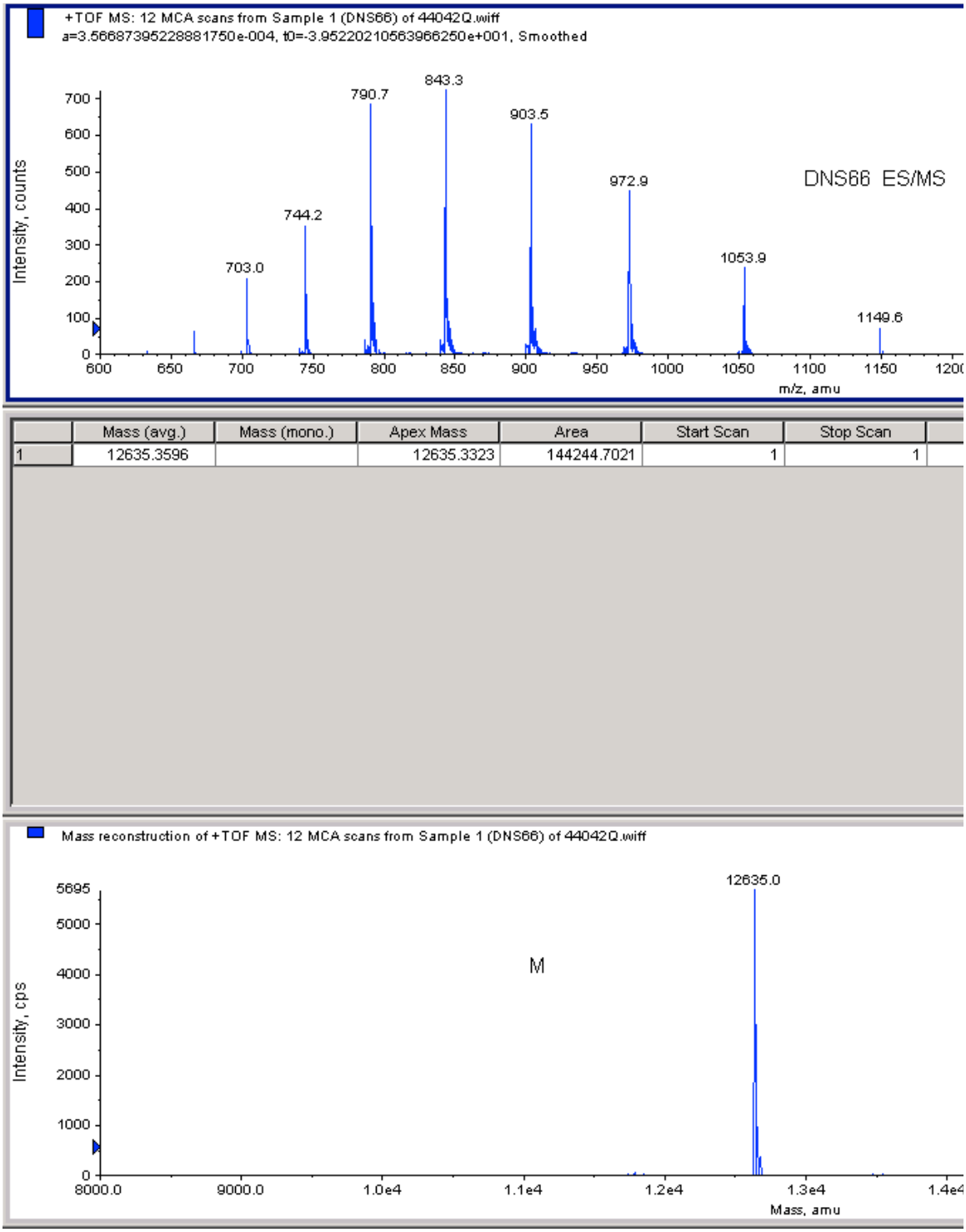

Figure 3.2: Representative electrospray ionization mass spectrum of Dns-labeled cytochrome cb_{562} (variant D66C).

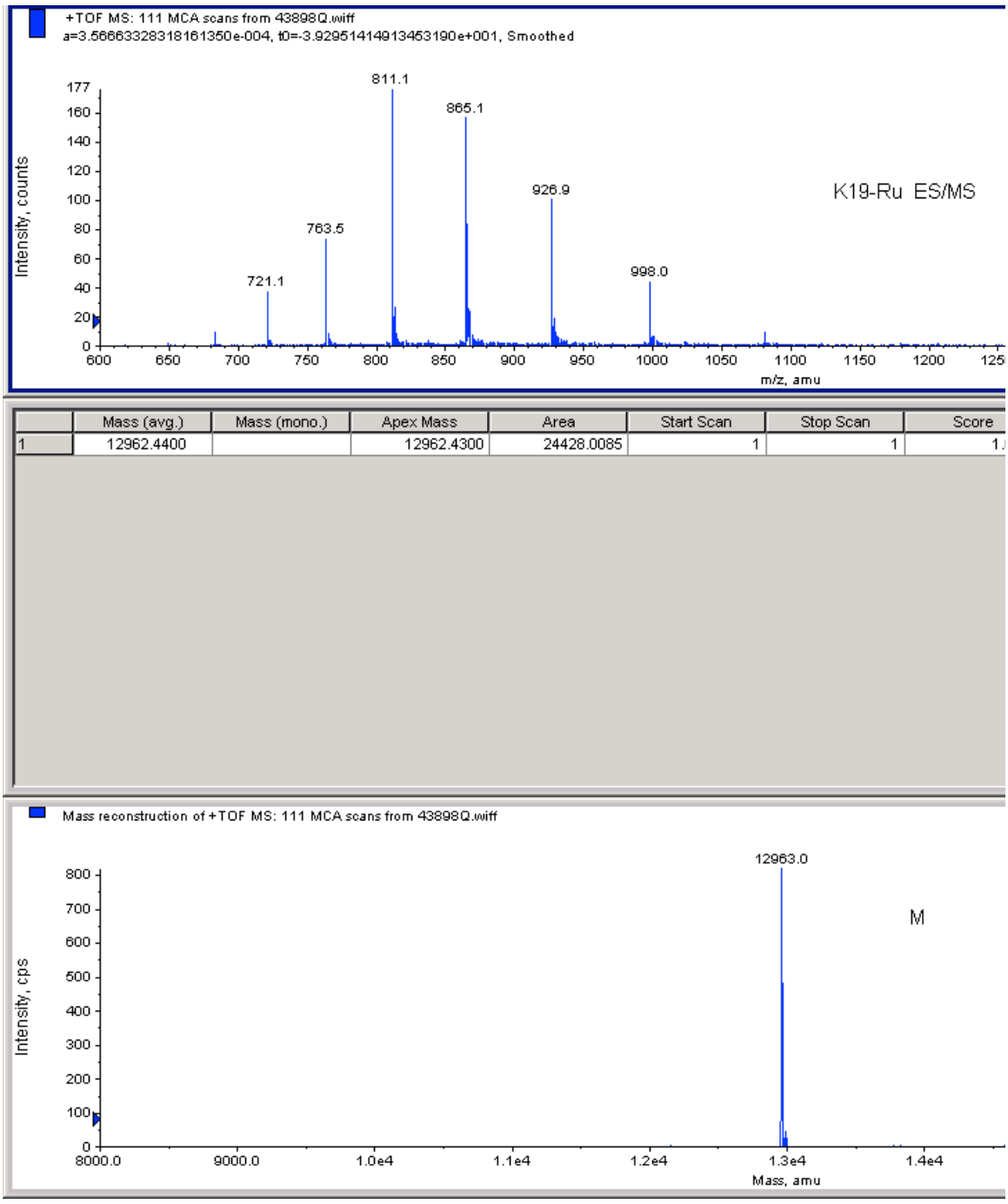

Figure 3.3: Representative electrospray ionization mass spectrum of Ru-labeled cytochrome cb_{562} (variant K19C).

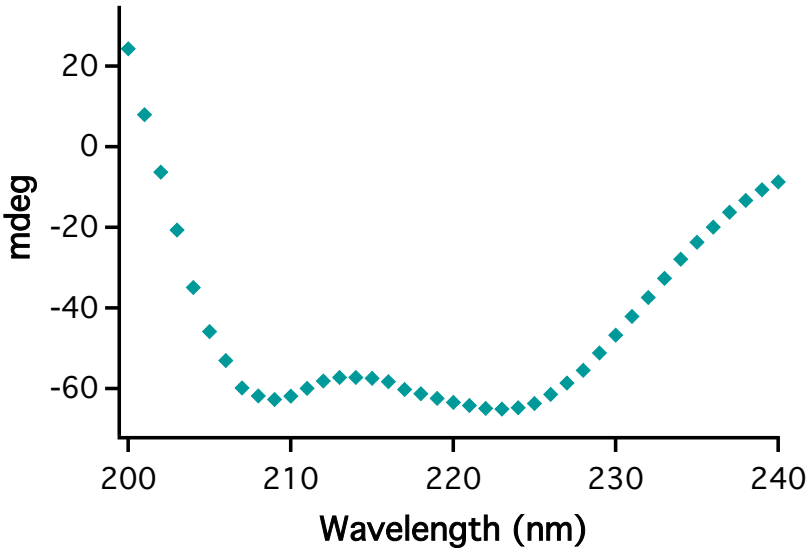

Figure 3.4: Representative circular dichroism spectrum of Dns-labeled cytochrome cb_{562} (variant E92C).

CD spectroscopy provides information about the secondary structure and helical content of a protein. The circular dichroism spectrum of cytochrome cb_{562} is shown in Figure 3.4. The two minima around 222 and 209 nm are indicative of α-helical structure. As the protein unfolds, the signal at 222 nm decreases until it reaches baseline (Figure 3.5).

UV-visible absorption spectroscopy probes the heme environment. As the protein unfolds, the heme Soret absorbance maximum red shifts. In the unfolding curve in Figure 3.6, the ratio of absorbance at the wavelengths that correspond to the maxima for the unfolded and folded protein are plotted.

The steep transition in the denaturation curves is indicative of a cooperative transition (Figure 3.7). The midpoint concentration of denaturant ($[\text{Gdn}]_{1/2}$) and a term representing the change in solvent accessibility upon unfolding (m) are

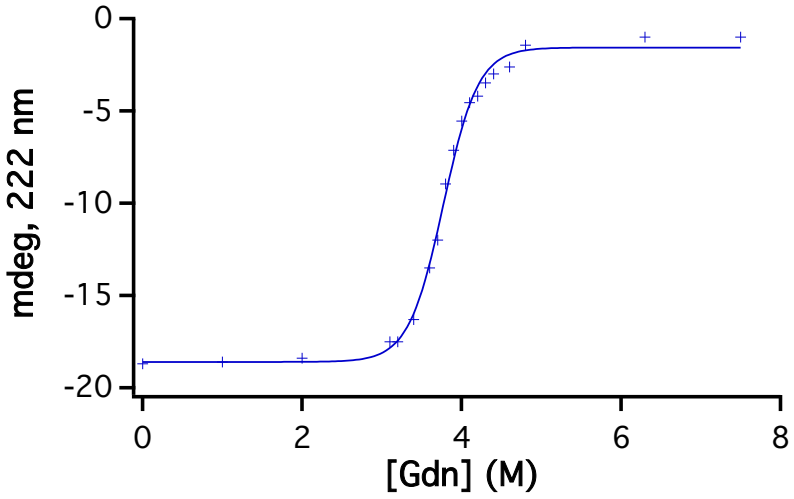

Figure 3.5: Chemical denaturation curve of Dns-labeled cytochrome cb_{562} probed by CD spectroscopy at pH 5 (variant E92C).

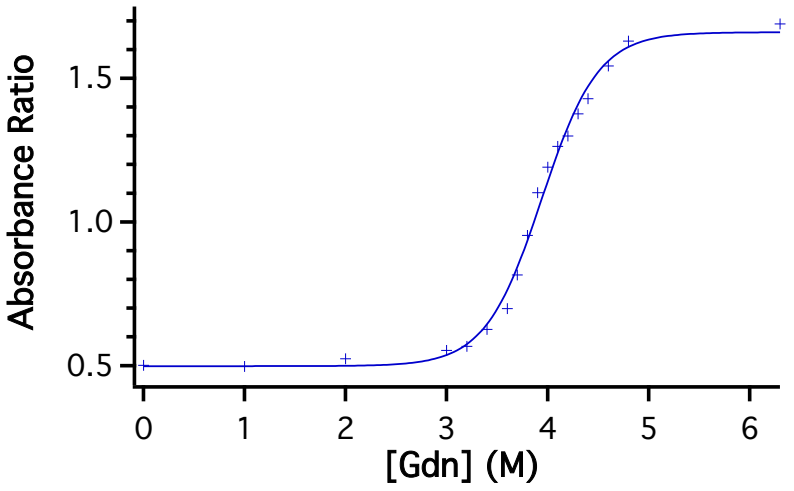

Figure 3.6: Chemical denaturation curve of Dns-labeled cytochrome cb_{562} probed by UV-visible absorption spectroscopy at pH 5 (variant E92C). The ratio of absorbance at 402 nm to 415 nm is plotted.

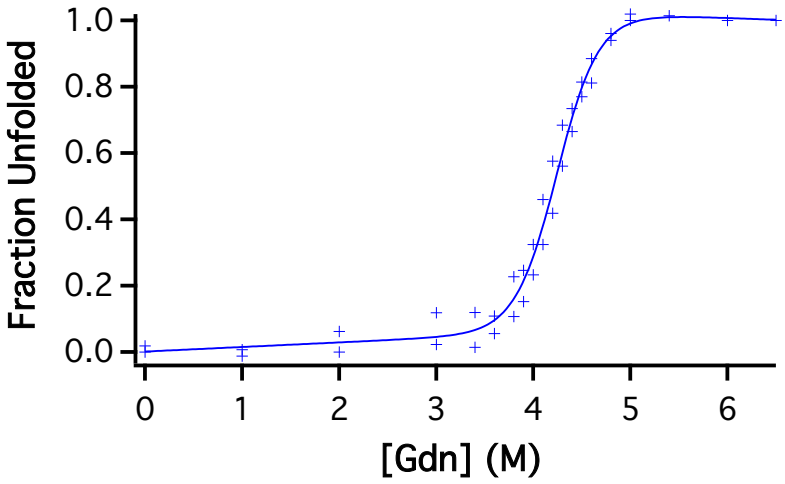

Figure 3.7: Chemical denaturation curve of Dns-labeled cytochrome cb_{562} (variant K19C, CD and UV-visible data).

Table 3.1: Denaturation parameters for Dns-labeled cytochrome cb_{562} variants.

Variant	$[\text{Gdn}]_{1/2}$ (M)	m (kJ/mol/M)	$-\Delta G$ (kJ/mol)
WT [36]	4.2	10	42
Dns19	4.2	11	46
Dns66	4.4	10	44
Dns92	3.8/3.9	12/9	40

extracted by fitting denaturant unfolding curves to a two-state model [59]. The free energy of folding at 0 M Gdn is the product of $[\text{Gdn}]_{1/2}$ and -m. $[\text{Gdn}]_{1/2}$ values for the labeled variants (Table 3.1) are comparable to the wild-type protein (4.2 M) [36], suggesting minimal perturbation of the structure and stability from incorporation of the cysteine mutation and Dns fluorophore. The Dns92 variant is slightly destabilized.

We have also measured the stability of cytochrome cb_{562} to thermal denaturation by CD spectroscopy (Figure 3.8). Cytochrome cb_{562} is exceptionally stable and contains significant secondary structure at 90°C. The partial unfolding that is

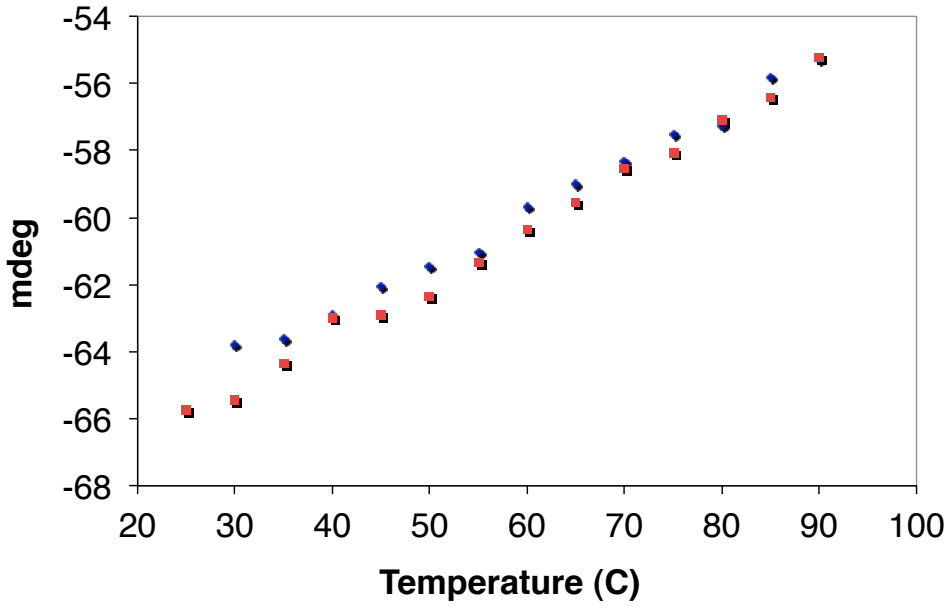

Figure 3.8: Temperature denaturation of Dns92-cytochrome cb_{562}. Unfolding is reversible; unfolding and refolding curves are shown in blue and red, respectively.

observed is not cooperative. Thus, heat-induced denaturation would be ineffective for studying the refolding of cytochrome cb_{562}.

3.4 Steady-State trFRET Measurements

We employed trFRET to characterize the conformational populations of Dns-labeled cytochrome cb_{562} variants upon denaturation with Gdn. The Gdn-dependence of the normalized integrated fluorescence intensity (Figure 3.9), comparable to the CD and absorption denaturation curves, is indicative of cooperative unfolding.

Distance distributions were extracted from fluorescence decays (Figures 3.10–3.12). At 0 M Gdn, Dns–heme distances are consistent with the crystal structure distances in the folded wild-type protein (Table 3.2); the most probable distances

Table 3.2: Donor-acceptor distances in angstroms for Dns-labeled cytochrome cb_{562} variants.

Variant	Distance C_γ-Fe*	r_{mode} †	r_{mean} †	Random coil distance‡	$r_{mode,\ 6\ M\ Gdn}$ †
19	19.8	22	23	96	≥50
66	14.9	19	21	62	≥50
92	16.3	20	20	28	22

∗Measured from the crystal structure (PDB 2BC5).
†Extracted from the Dns-heme distance distributions at pH 4.
‡Predicted by calculation [60].

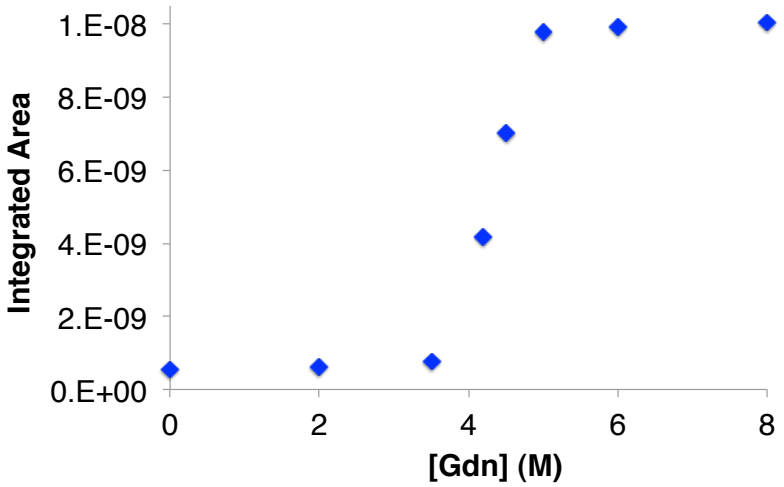

Figure 3.9: Denaturant-dependence of integrated normalized Dns fluorescence decays for Dns19-cytochrome cb_{562}.

extracted from the fluorescence decays (r_{mode}) are slightly longer (2–4 Å) than the Cγ-Fe distances in the crystal structure, likely owing to the length of linker between the Dns fluorophore and the Cγ atom. Denaturation with Gdn slowed the fluorescence decay in all Dns variants, consistent with an increase in the average Dns–heme distance. Dns fluorescence decays could not be fit to single exponential functions as distance components exist with various probabilities, consistent with the presence of multiple conformations.

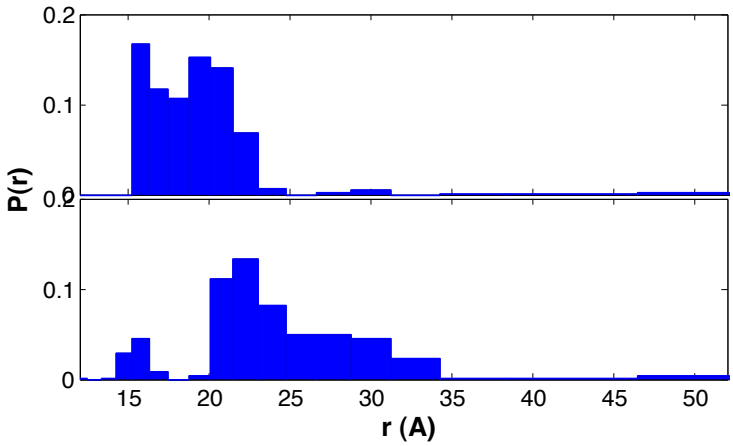

Figure 3.10: Dns92–heme distance distributions for native (upper) and denatured (8 M Gdn, lower) cytochrome cb_{562} at pH 5 by trFRET.

A random coil is characterized by a statistical distribution of conformations that lack well-defined structure. The most probable Dns–heme distances for the extended states roughly correspond to the calculated random coil distances (Table 3.2). The extended conformations of Dns92 (Figure 3.10) are distributed around this random coil distance. The predicted random coil distances for mutants Dns19 and Dns66 are beyond the resolution of the measurement ($>$ 52 Å); however, notable populations at $\geq$50 Å are consistent with the predicted distances (Figures 3.11–3.12). Significantly, the extended structures are more populated for mutant K19C, which has a longer calculated random coil distance. In contrast, the unfolded conformations of cytochrome *c′* do not resemble a random coil; trFRET measurements of Trp32–heme distances are similar to Trp72-heme distances, despite the significantly greater distance between Trp32 and the heme in the polypeptide chain [28].

Dns92, located on helix 4, probes formation of the secondary structure of

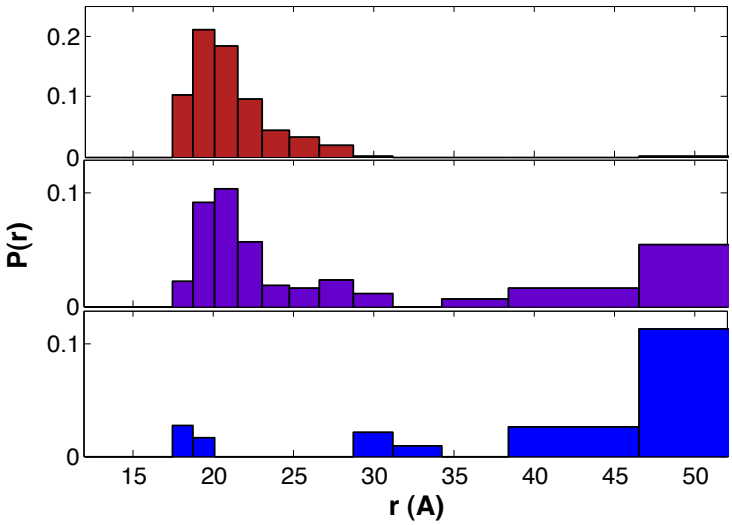

Figure 3.11: Dns66–heme distance distributions of cytochrome cb_{562} in 0 M Gdn (red), 4.2 M Gdn (violet), and 6 M Gdn (blue) at pH 4 by trFRET.

this helix. Residue 92 is located near the heme on the polypeptide chain, so only modest broadening of the D-A distribution is observed upon denaturation (Figure 3.10). Despite the overlap of states in the unfolded ensemble, there is an appreciable occupation of extended states (>25 Å). The compact unfolded conformations could consist of bent non-helical polypeptide conformations that bring the Dns fluorophore and heme close together, as well as native-like contacts; heme misligation may increase the occupation of this compact population (see section 3.5).

Variant Dns66 probes tertiary contacts between helix 3 and the heme, attached to helix 4 (Figure 3.11). The unfolded state ensemble consists predominantly of extended structures (94% at $\geq$ 30 Å, 86% at $\geq$ 40 Å). At the denaturation midpoint ([Gdn] = 4.2 M), the protein ensemble equally resembles the folded and unfolded protein with 47% extended conformations ($>$ 30Å).

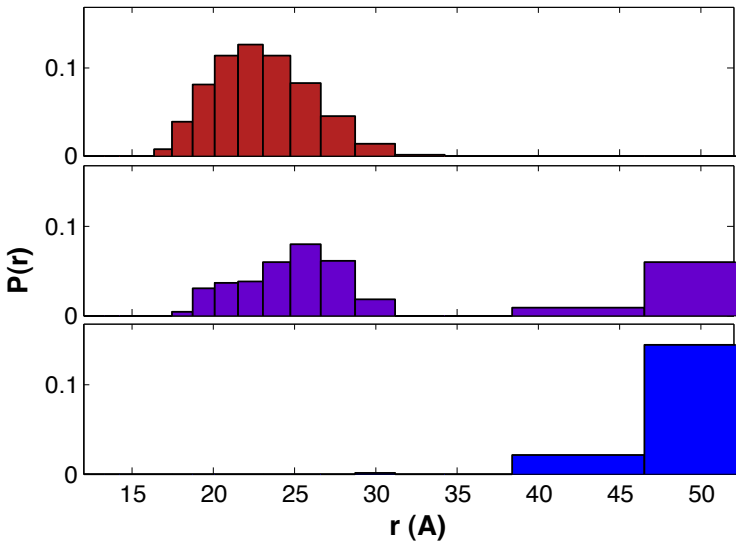

Figure 3.12: Dns19–heme distance distributions of cytochrome cb_{562} in 0 M Gdn (red), 4.2 M Gdn (violet), and 6 M Gdn (blue) at pH 4 by trFRET.

Variant Dns19 probes tertiary contacts between helix 1 and the heme (Figure 3.12). The unfolded state ensemble is populated exclusively with extended structures. At the denaturation midpoint, two populations with distributions that resemble the native and unfolded states are observed.

3.5 Heme Environment[1]

The heme coordination of cytochrome cb_{562} is shown in Figure 3.13. When the protein is folded, the heme is axially ligated to the sidechains of Met7 and His102. Ferricytochrome (Fe^{3+}) cb_{562} is unfolded by Gdn with a denaturation midpoint of 4.2 M Gdn. Changes in the heme environment with different solvent conditions were probed by UV-visible spectroscopy (Figure 3.14).

Ferricytochrome cb_{562} has a Soret maximum of 415 nm. When unfolded in 6 M Gdn at pH 4, the Soret maximum blue shifts to 400 nm as the heme transitions from low spin to high spin, suggesting that the Met7 sulfur ligation is replaced by a water molecule [37]. As the pH of the unfolded protein is increased, the Soret peak red shifts (λmax = 412 nm). The shift and pH range are consistent with heme ligation by a histidine imidazole. This observation suggests that in denatured cyt cb_{562}, the nonnative histidine ligand (His63) coordinates to the heme at pH $\geq$ 5.

Conformational populations of Dns66-cytochrome cb_{562} were determined by trFRET measurements at variable Gdn concentrations and pH (Figures 3.15). Extended conformations with >50-Å distances comprise a significant population of denatured proteins (6 M Gdn). At pH 4, there are few compact states. With increasing pH, however, the proportion of the compact population increases, consistent with a decrease in the configurational freedom and the number of states available to the unfolded protein upon formation of a 36-residue loop by His63 ligation to the heme.

[1] Adapted from Bouley Ford, N.; Shin, D.W.; Gray, H.; Winkler, J. "Tertiary Contact Dynamics in Unfolded Cytochrome cb_{562}" *Submitted*.

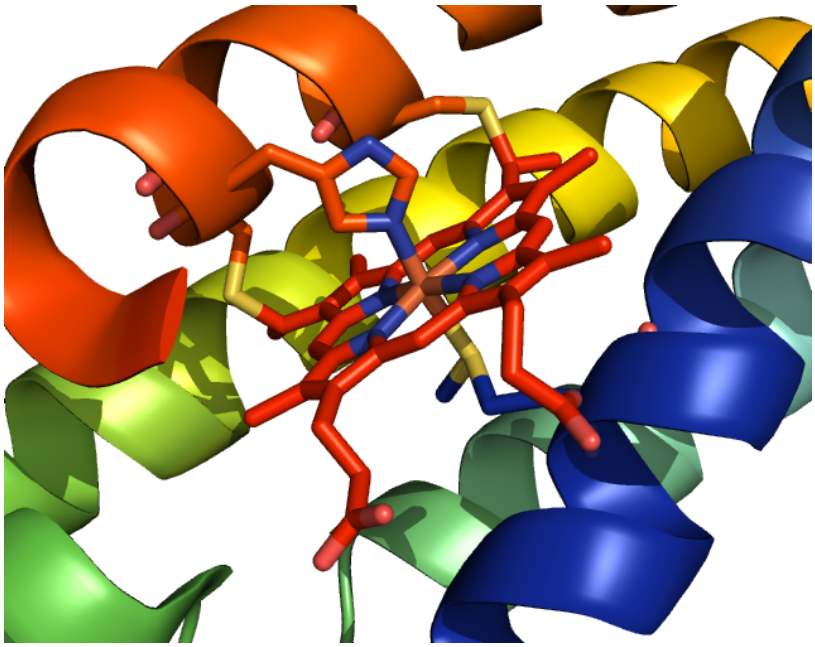

Figure 3.13: Heme environment of cytochrome cb_{562} (PDB 2BC5). The heme (red) is attached covalently to the fourth helix (orange) and ligated to Met 7 on the first helix (blue) and His102 on the fourth helix.

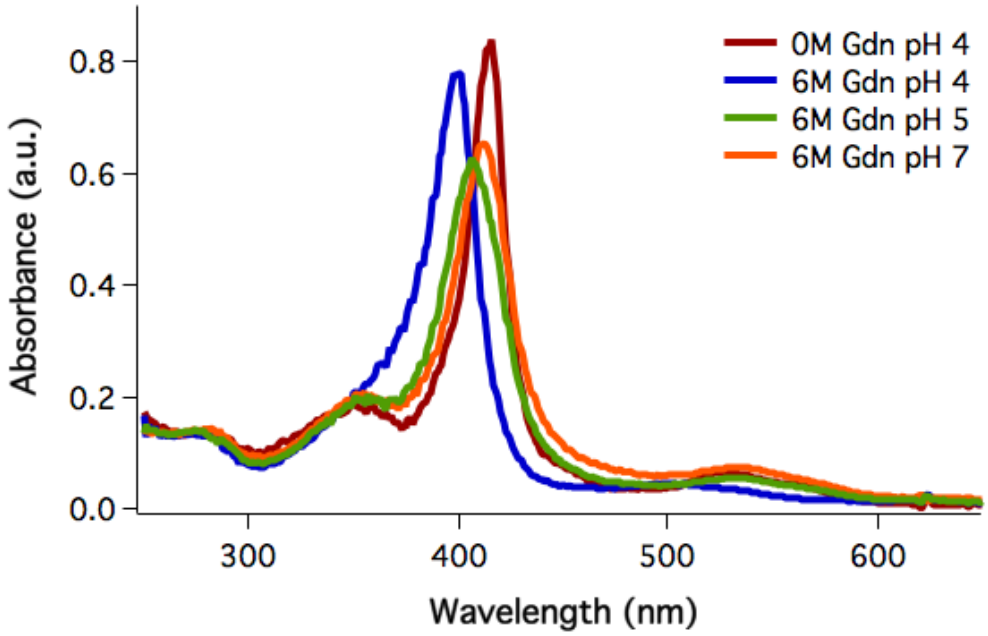

Figure 3.14: UV-visible spectra of cytochrome cb_{562}, variant K32C. The heme is ligated by Met7 and His102 when folded (λ_{max} = 415 nm). When the ferricytochrome is denatured by 6 M Gdn, Met7 is replaced by water or His63 (λ_{max} = 400 nm at pH 4, 406 nm at pH 5, 412 nm at pH 7).

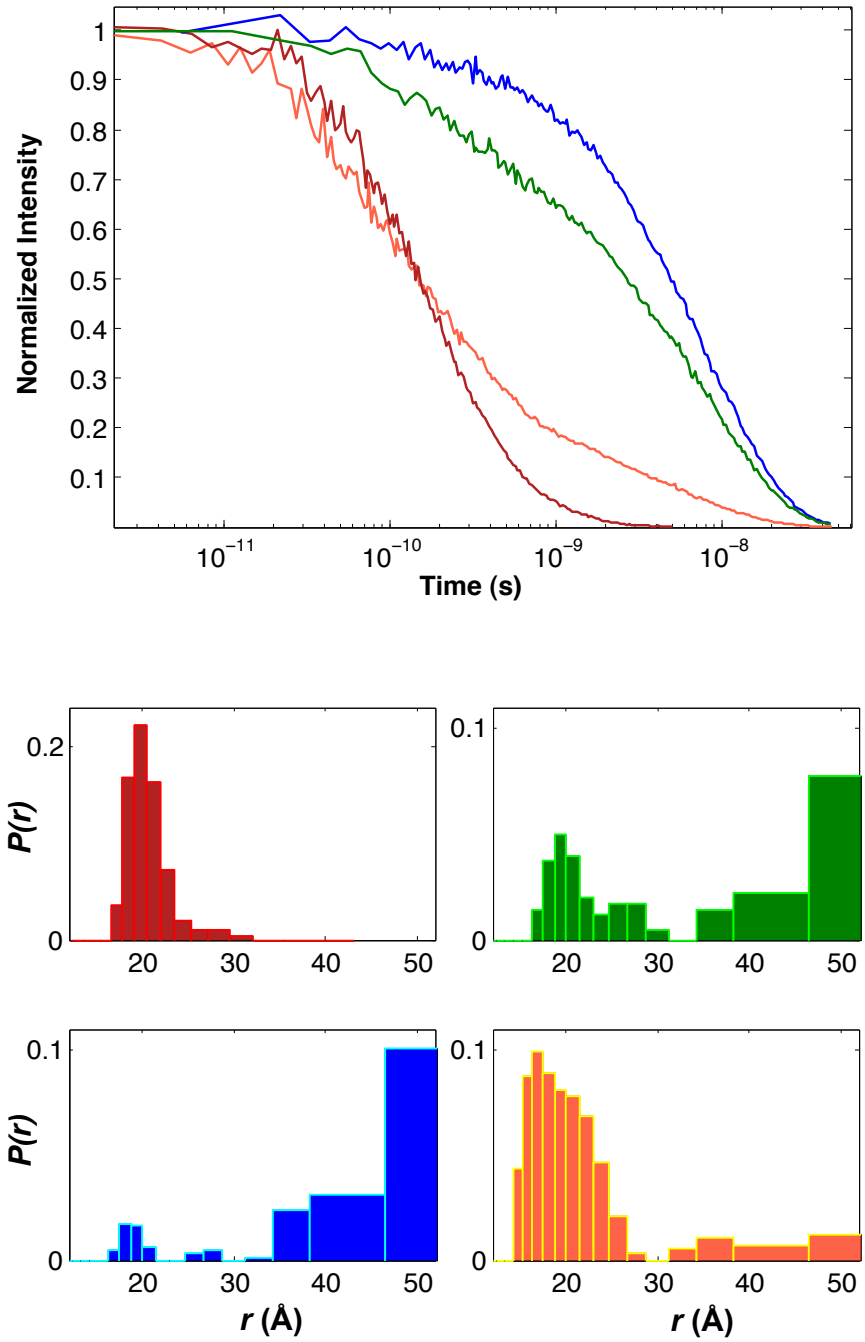

Figure 3.15: Fluorescence decays and corresponding distance distributions of Dns66-cytochrome cb_{562}: folded at pH 4 (red) and unfolded in 6 M Gdn at pH 4 (blue), pH 5 (green), and pH 7 (orange).

3.6 Conclusions

Cytochrome cb_{562} is highly stable to chemical (Gdn) and thermal denaturation, exhibiting incomplete unfolding at 90°C. The protein exhibits cooperative unfolding in equilibrium conditions ($D_{50\%}$ = 4.2 M Gdn) by circular dichroism, UV-visible absorbance, and fluorescence spectroscopy.

The properties of the unfolded state ensemble strongly influence the folding mechanism. At pH 4, the unfolded state, probed by trFRET with a Dns fluorophore attached to K19C, D66C, or E92C, resembles a random coil. However, pH studies of unfolded cytochrome cb_{562} by UV-visible absorption spectroscopy indicate that the heme likely coordinates to the nonnative ligand His63 at pH > 4. Misligated conformations, probed by trFRET measurements of Dns66, are significantly more compact.

3.7 Acknowledgments

Parts of this work were carried out in collaboration with Caltech undergraduates Yuehan Huang and Dong Woo Shin.

Chapter 4

Tertiary Contact Dynamics in Unfolded Cytochrome cb_{562} by Electron Transfer[1]

4.1 Introduction

When studying the dynamics of unfolded proteins, frustration manifests as internal friction, slowing intrachain diffusion in regions of the polypeptide. As formation of tertiary contacts is a necessary step in folding, the speed limit is set by the conformational dynamics in the unfolded state [14, 15, 16, 17]. Statistical mechanical models suggest that the rates of tertiary contact formation in polymers are determined by the intrachain diffusion coefficient and chain length [61, 62]. Approximate analytical solutions to the Smoluchowski equation for idealized proteins predict that contact rate constants will show a power-law distance dependence on n ($n^{-1.5}$), where n is the number of peptide bonds between contacts [61]. A model that takes into account the loop size formation probability predicts a

[1] Adapted from Bouley Ford, N.; Shin, D.W.; Gray, H.; Winkler, J. "Tertiary Contact Dynamics in Unfolded Cytochrome cb_{562}" *Submitted*.

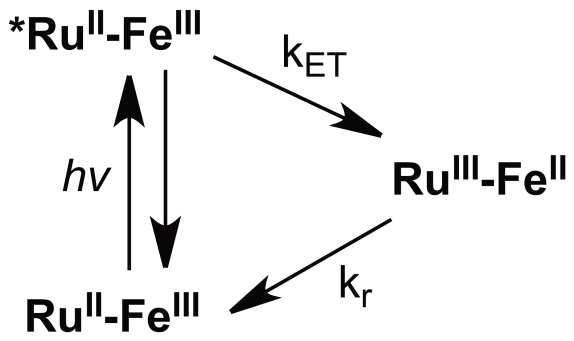

Figure 4.1: Luminescence of photoexcited $*Ru^{II}$ is quenched by electron transfer upon contact with the ferriheme.

maximum rate at n = 10 with an $n^{-3.2}$ dependence for larger loops and slower rates for small loops, owing to increased stiffness [62]. Analysis of contact quenching experiments probing intrachain diffusion in unstructured polypeptides revealed a power-law dependence for large loops with an n-independent asymptote for small loops [63, 64].

In this chapter, we describe the contact dynamics of denatured cytochrome cb_{562}, a minimally frustrated protein. We examine the formation of transient contacts between the heme and $[Ru(bpy)_2(IA\text{-}phen)]^{2+}$ complexes (Figure 1.7) that have been covalently attached at various positions in the protein chain. Rates of intrachain diffusion in chemically denatured cytochrome cb_{562} molecules have been obtained from analysis of time-resolved luminescence measurements, thereby allowing us to assess the degree of frustration in the energy landscape of this four-helix bundle. We also have found that forming one tertiary contact can accelerate formation of subsequent contacts.

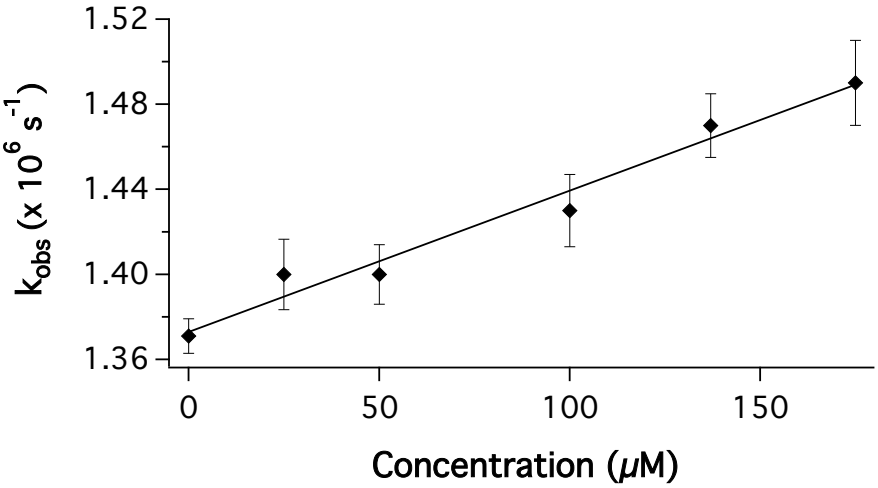

Figure 4.2: Stern-Volmer plot for the quenching of photoexcited $Ru(bpy)_3^{2+}$ by cytochrome cb_{562} (Fe^{3+}) in 6 M Gdn at pH 4. The data fit to a k_Q of 6.6($\pm$0.5) x 10^8 $M^{-1}s^{-1}$.

4.2 Contact Quenching

Following photoexcitation, $*[Ru(bpy)_2(IA\text{-}phen)]^{2+}$ luminesces with an unquenched lifetime of $\sim$1 μs (k_{label} = 1.10($\pm$0.04) x 10^6 s^{-1} at pH 4–7 with 6–8M Gdn). Contact with a ferriheme quenches the $*Ru^{II}$ state via electron transfer (Figure 4.1). We have performed quenching studies to determine the rate constant for the bimolecular electron transfer reaction using $Ru(bpy)_3^{2+}$ as a model for the Ru-label $[Ru(bpy)_2(IA\text{-}phen)]^{2+}$. Quenching was examined at several protein concentrations in 6 M Gdn (Figure 4.2). We measure a quenching rate constant (k_Q) of 6.6($\pm$0.5) x 10^8 $M^{-1}s^{-1}$. Correcting for the viscosity of 6 M Gdn [65] gives k_Q = 1.2 x 10^9 $M^{-1}s^{-1}$, consistent with a reaction near the diffusion-controlled limit.

A D L E D N M E T L N D N L K V I E K A D N A A Q V K D A L 30
T K M R A A A L D A Q K A T P P K L E D K S P D S P E M W D 60
F R *H* G F D I L V G Q I D D A L K L A N E G K V K E A Q A A 90
A E Q L K T T C N A C H Q K Y R 106

Figure 4.3: Sequence of cytochrome cb_{562}. Cysteines by which the porphyrin is attached covalently to the protein backbone are shown in red. Heme ligands are shown in orange with the non-native ligand His63 italicized. Residues that are mutated to cysteine in single-mutation variants are shown in blue.

4.3 Ruthenium-Labeled Variants

Residues labeled with $[Ru(bpy)_2(IA\text{-}phen)]^{2+}$ in single-cysteine variants are highlighted in the sequence of cytochrome cb_{562} in Figure 4.3. The circular dichroism (Figure 4.4) and heme Soret absorption spectra of labeled proteins are virtually identical with those of the unlabeled wild-type, confirming that the mutations and Ru-labeling do not disrupt the protein secondary structure or heme environment.

4.4 Transient Contact Formation Rates

The formation of transient loops in the polypeptide chain of cytochrome cb_{562} was probed by monitoring electron transfer between the heme and $[Ru(bpy)_2(IA\text{-}phen)]^{2+}$ complexes attached at various positions on the polypeptide chain. The porphyrin vinyl groups of the heme are covalently attached to the protein backbone of this 106-residue protein via two mutant cysteine residues, R98C and Y101C, in a CXXCH cyt *c*-type motif [37]. Data were collected at pH 4 to inhibit His63 ligation of the heme and allow the denatured proteins a full range of motion.

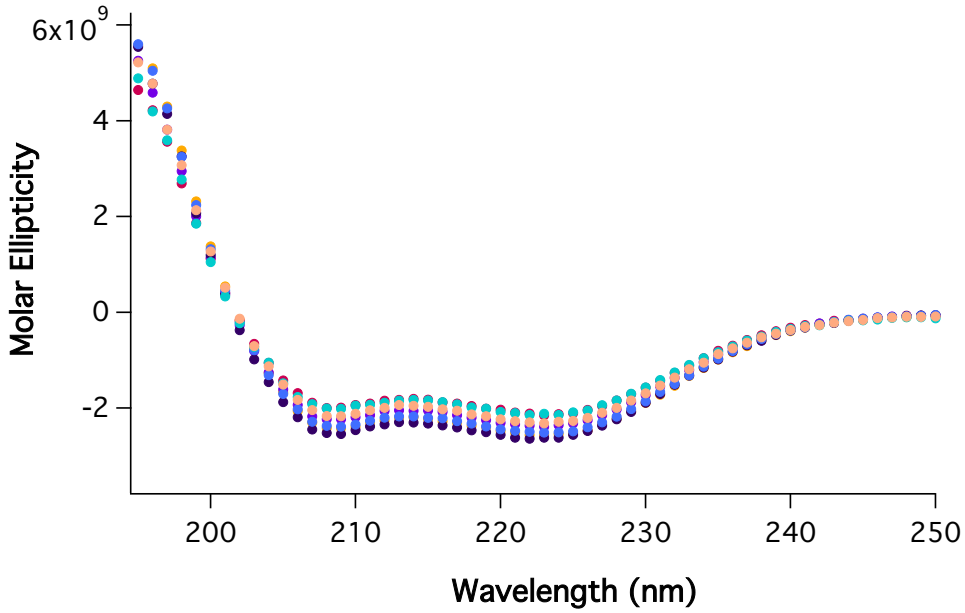

Figure 4.4: Circular dichroism spectra of cytochrome cb_{562} Ru-variants at pH 7: Ru19 (yellow), Ru32 (red), Ru51 (purple), Ru66 (deep purple), Ru77 (blue), Ru83 (teal), and Ru92 (coral).

As bimolecular electron transfer is diffusion-limited, the rate constant for the labeled proteins ($k_{ET} = k_{obs} - k_{label}$) corresponds to the specific rate of transient contact formation.

Luminescence decays for Ru-variants in 6 M Gdn and the corresponding $1/k_{ET}$ values are shown in Figures 4.5 and 4.6. The degree of quenching correlates with the efficiency of the electron transfer reaction. The luminescence decays exhibit single-exponential kinetics, consistent with interconversion among protein conformations that is rapid compared to luminescence decay. Luminescence decay rate constants (k_{obs}) for the variants range from 1.53($\pm$0.03) x 10^6 s^{-1} for Ru51 to 9.4($\pm$0.6) x 10^6 s^{-1} for Ru92 (Table 4.1). Luminescence decays for Ru19 and Ru32 with the greatest Ru-heme distance are not quenched. For these variants, electron

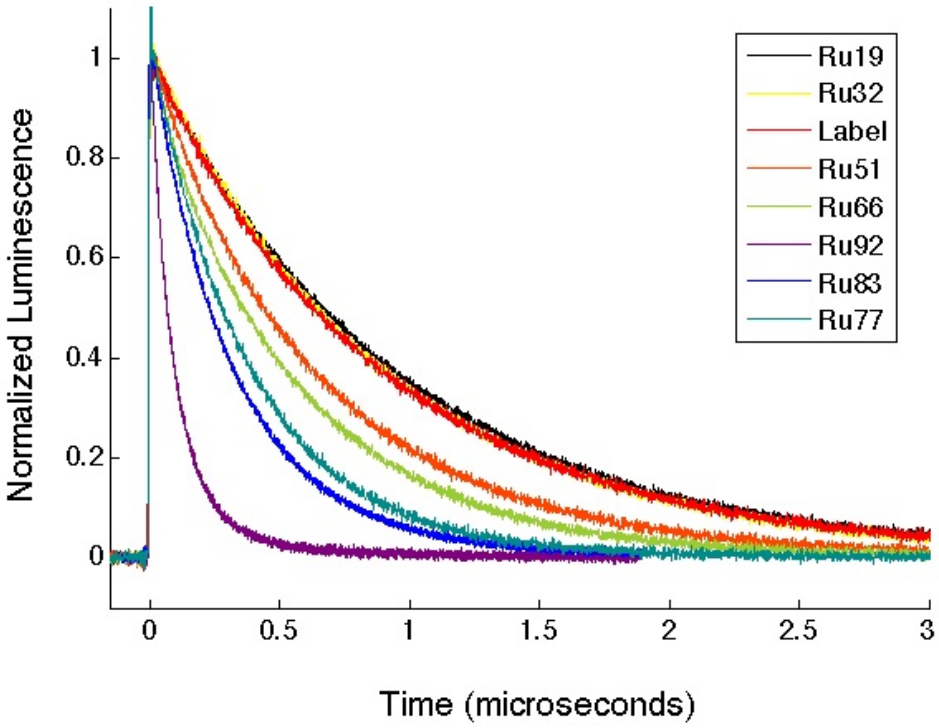

Figure 4.5: Luminescence decays of $^*Ru^{II}$ in the Ru-labeled cytochrome cb_{562} variants and the unquenched free label at pH 4 in 6 M Gdn.

transfer is not competitive with excited state deactivation, indicative of a contact time greater than 20 μs (the standard deviation for k_{ET}^{-1}).

We observe a Gdn-dependence of contact formation rate constants (k_{ET}), as expected for a diffusion-limited process. For example, k_{ET} for Ru77 is 1.37($\pm$0.07) x 10^6 s^{-1} at 6 M Gdn and 9.1($\pm$0.5) x 10^5 s^{-1} at 8 M Gdn. A viscosity dependence also has been observed for intrachain diffusion in denatured cytochrome *c* and in Gly-Ser-repeat polypeptides [15, 64]. Prior work on end-to-end contact kinetics in Gly-Ser-repeat polypeptides found a linear dependence of ln(k) on [Gdn] [23]. If we assume similar behavior in cytochrome cb_{562}, then our two data points are consistent with a slope of $-\Delta\ln(k)/\Delta[Gdn] = 0.2$ M^{-1}, comparable to the values

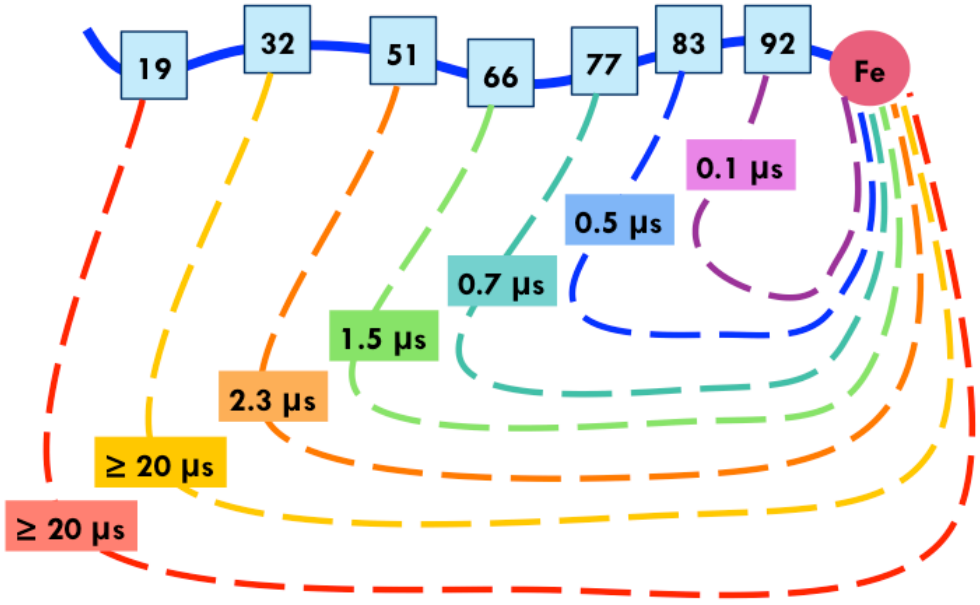

Figure 4.6: Schematic of the contact formation time constants for denatured Ru-cytochrome cb_{562} variants, which show a distance dependence.

Table 4.1: Fitted rate constants for luminescence data collected at pH 4 in 6 M Gdn and the corresponding contact formation rate constants.

Variant	Distance (n)	k_{obs} (s^{-1})	k_{ET} (s^{-1})
Free label	N/A	$1.09(\pm0.03) \times 10^6$	N/A
Ru19	79	$1.06(\pm0.02) \times 10^6$	Not observed
Ru32	66	$1.14(\pm0.04) \times 10^6$	Not observed
Ru51	47	$1.53(\pm0.03) \times 10^6$	$4.4(\pm0.4) \times 10^5$
Ru66	32	$1.75(\pm0.04) \times 10^6$	$6.6(\pm0.5) \times 10^5$
Ru77	21	$2.46(\pm0.07) \times 10^6$	$1.37(\pm0.07) \times 10^6$
Ru83	15	$3.1(\pm0.2) \times 10^6$	$2.0(\pm0.2) \times 10^6$
Ru92	6	$9.4(\pm0.6) \times 10^6$	$8.3(\pm0.6) \times 10^6$

extracted from Gly-Ser-repeat polypeptides (0.19–0.23 M^{-1}) [23]. Extrapolating to 0 M Gdn solution suggests that the Ru77-heme contact rate constant would be 5 x 10^6 s^{-1} in buffer solution without Gdn. Applying the same extrapolation to our largest measured rate constant (Ru92, 10^7 s^{-1} in 6 M Gdn) leads to an estimated contact time of ~30-ns for Ru92-heme in water.

The contact formation rate constants display a power-law dependence on *n*, the number of peptide bonds between the Ru-label and the heme (Figure 4.7). A fit to the function $k_{ET} = A*n^y$ gives the following parameters: A = 1.20($\pm$ 0.09) x 10^8 s^{-1} and y = -1.49 $\pm$ 0.04. The random polymer model by Szabo, Schulten, and Schulten predicts this trend ($k \propto n^{-1.5}$) for purely entropy-controlled intrachain diffusion in ideal, freely-jointed Gaussian chains [61]. In contrast, end-to-end contact measurements of unstructured Gly-Ser-repeat polypeptides of variable length in water and 8 M Gdn exhibit power-law slopes of -1.7 and -1.8, respectively, which are consistent with the Gaussian-chain model when excluded-volume effects are included [64]. Importantly, our data indicate that cytochrome cb_{562} behaves as a random coil when denatured and that excluded volume corrections to our intrachain contact rates are not needed.

The adherence of the data ($n \geq 6$) to the random polymer model provides strong support for the proposal that cytochrome cb_{562} (helix 2/3 loop to the heme) is minimally frustrated. Deviations from the power-law function would be expected for frustrated regions that exhibit increased internal friction. And indeed, deviations were found for α-synuclein, the intrinsically disordered polypeptide

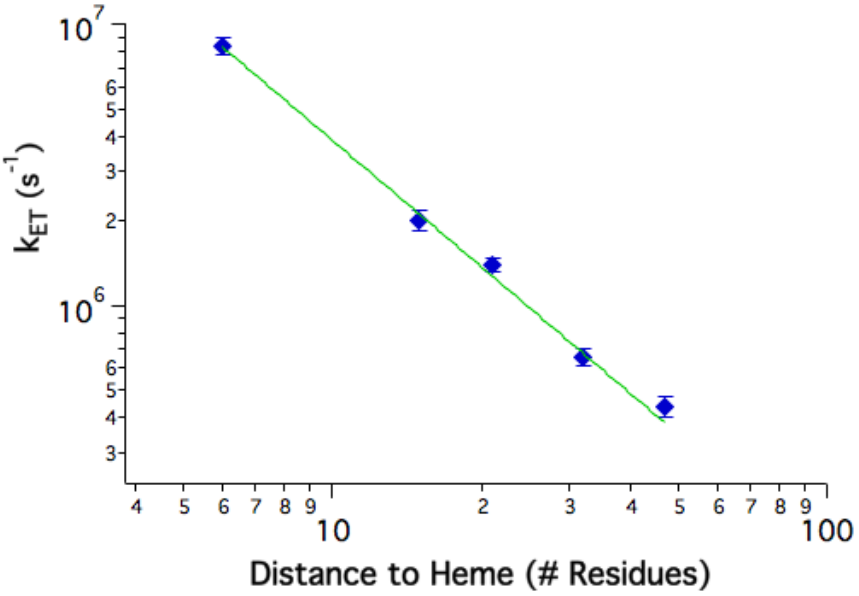

Figure 4.7: Log/log plot of contact formation rates versus number of peptide bonds between transient intramolecular contacts (Table 4.1). Ru-variants are denatured in 6 M Gdn at pH 4. The curve fits to a power function with a slope of -1.49 $\pm$ 0.04.

implicated in Parkinson's disease [19]. It will be of special interest to compare the unfolded dynamics of cytochrome cb_{562} to those of cytochrome *c'*, as the latter protein displays probe-dependent folding kinetics [28].

Our contact rate constants for the denatured cytochrome cb_{562} polypeptide are substantially smaller than those reported for Gly-Ser-repeat polypeptides in Gdn, a trend that also was found for the contact dynamics of α-synuclein in the absence of denaturants [64, 18]. For comparison, contact formation for Ru77 (n=21) in 8 M Gdn is five times slower than for the repeat polypeptides in 8 M Gdn [64]. The different probes of contact events (i.e., triplet-triplet energy transfer versus electron transfer) may contribute to the discrepancy. If the *Ru to heme electron transfer reaction is not strictly diffusion controlled, the observed contact times could be

systematically greater than those found using energy-transfer quenching.

Alternatively, the intrachain diffusion coefficient in cytochrome cb_{562} may be smaller than that for repeat polypeptides. Although Gdn viscosity effects, as we have noted, are similar for cytochrome cb_{562} and repeat polypeptides, the increased conformational freedom of glycine as compared to other amino acids could contribute to a somewhat increased diffusion coefficient in the synthetic peptides. The presence of glycine can only partially explain the difference however; polyserines diffuse two times slower than Gly-Ser polypeptides [64].

Dynamical drag introduced by segments of the polypeptide chain exterior to the loop also likely contributes to the slower contact rates found in the 106-residue protein [66]. In synthetic polypeptides, interior-to-end contact rates decrease by as much as 2.5-fold compared to end-to-end contacts [67]. In proteins, interior contact dynamics would be more affected by drag from the external protein chain than contacts toward the ends of the chain. This drag could account for the differences in the power-law dependence for interior loop formation in cytochrome cb_{562} (y = -1.5) compared to that found for end-to-end contact formation (y = -1.8) in synthetic peptides.

4.5 Dynamics within a Constrained Loop

At pH $>$ 4, Ru66 quenching kinetics probe contact formation within a 36-residue loop formed by the His63 misligation to the heme (Figure 4.8). Under these conditions, Ru66 cytochrome cb_{562} exhibits biexponential luminescence decay

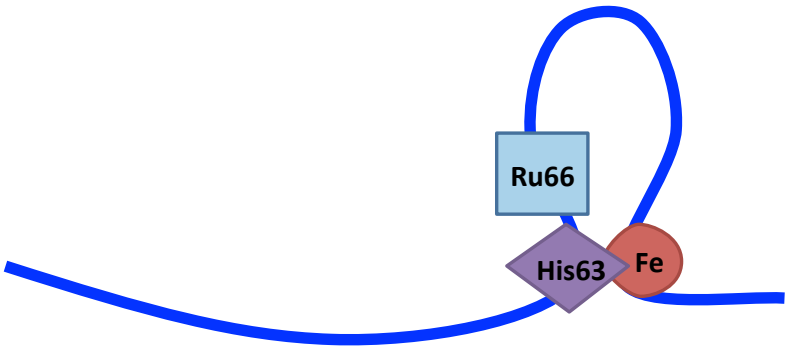

Figure 4.8: Schematic of long-lived loop formed by ferriheme misligation with His63 at pH > 4.

Table 4.2: pH dependence of contact formation rate constants for Ru66 in 6 M Gdn.

pH	First k_{ET} (s^{-1})	Second k_{ET} (s^{-1})
4	6.6 x 10^5 (100 %)	N/A
5	6 x 10^5 (60 %)	1.2 x 10^7 (40 %)
7	3 x 10^5 (17 %)	1.0 x 10^7 (83 %)

kinetics (Figure 4.9), consistent with two populations that are not exchanging on the microsecond timescale. The UV-visible spectrum of cytochrome cb_{562} suggests that there is an increase in the degree of His63 heme misligation from pH 5 to 7 as the Soret band further red shifts, concurrently with an increase in the compact conformational population observed in trFET measurements (see section 3.5). Over this pH range, the contribution of the second rate constant to the biexponential fit increases from 40% to 83% (Table 4.2 and Figure 4.10). It is likely, therefore, that the additional compact population is attributable to heme-misligated structures.

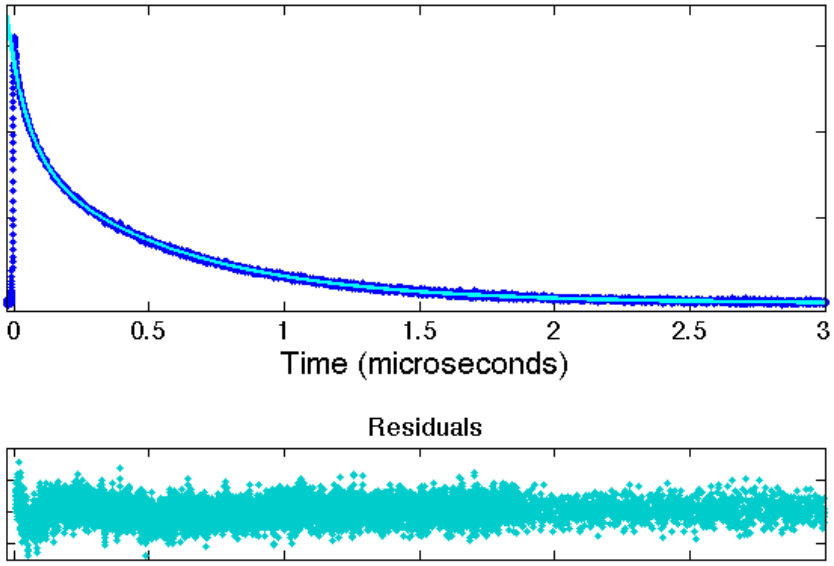

Figure 4.9: Biexponential fit of luminescence decay of Ru66 at pH 5.

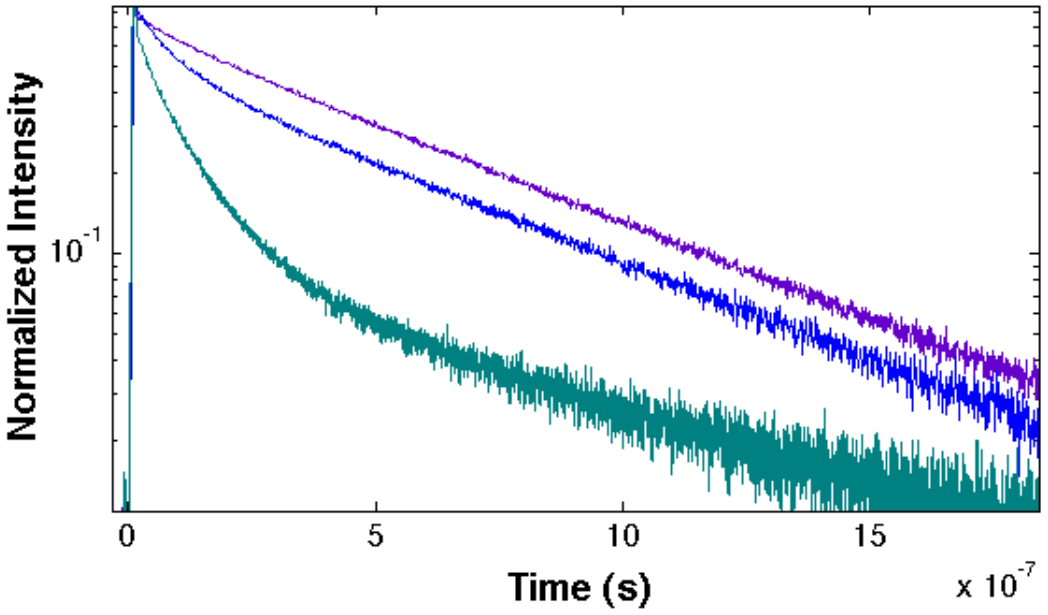

Figure 4.10: Luminescence decays of Ru66 at pH 4 (purple), pH 5 (blue), and pH 7 (teal). At pH 5 and 7, biexponential quenching by two distinct populations, the freely diffusing and misligated forms, is observed.

The 36-residue static loop formed by misligation serves as a model for early formation of tertiary contacts in protein folding. Ru66-heme contacts form more rapidly for misligated populations ($k_{ET} \approx 10^7$ s^{-1}) than in the random coil ($k_{ET} \approx 10^{5.8}$ s^{-1}). The heme-His63 contact acts as a nucleation site that facilitates interactions with nearby Ru66, with formation of this new contact approaching the folding speed limit. In contrast to Ru66, two variants between His63 and the heme, Ru77 and Ru83, do not show enhanced electron transfer in the misligated population; single exponential luminescence fits at pH 7 correspond to pH 4 fits (k_{ET} = 1.4 x 10^6 s^{-1} for Ru77 at both pH 4 and 7; k_{ET} = 2.0 x 10^6 s^{-1} and 8.4 x 10^5 s^{-1} for Ru83 at pH 4 and 7, respectively). Although the His63-heme contact speeds formation of a new contact with nearby Ru66, it does not assist residues further away in the loop. This behavior is consistent with a cooperative folding process. If a native tertiary contact is formed early in the folding pathway, it can limit the conformational search for other tertiary contacts by effectively decreasing n for nearby residues.

We observe a divergence from the power-law behavior when we include the misligated populations of Ru66 (Figure 4.11). Contact formation may be slowed due to decreased access to the heme when ligated by a histidine side chain instead of a water molecule; this rate constant could be a lower limit for loops of comparable size. Kiefhaber and coworkers found that the increased rigidity of smaller loops produces a limiting contact rate constant (k_0): $k = (1/k_0 + 1/(A*n^y))^{-1}$ [64]. The influence of the rate-limiting factor was less pronounced in Gdn than

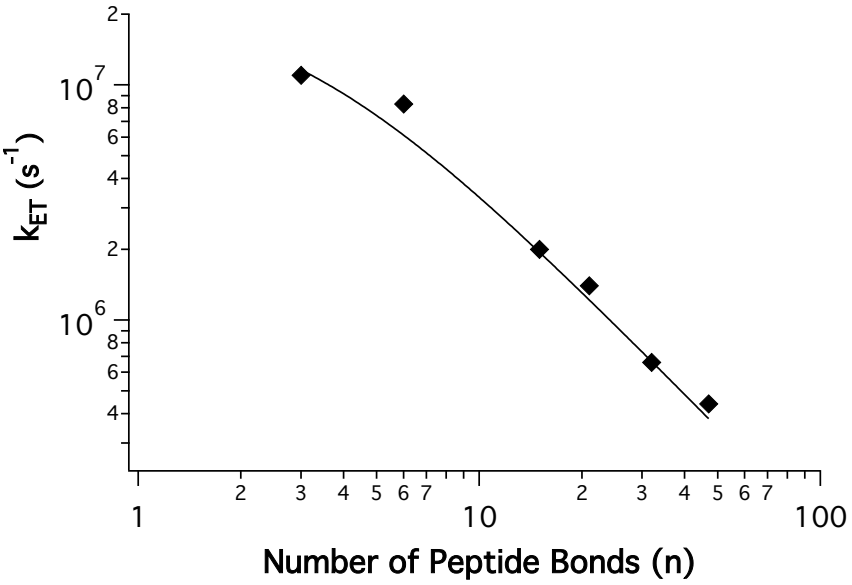

Figure 4.11: Fit of contact rate constants to $k = (1/k_0 + 1/(A*n^y))^{-1}$ with the following constraints: $A = 1.2 \times 10^8$ s^{-1} and $y = -1.5$. Rate constants reach an upper limit of $k_0 = 2.3 \times 10^7$ s^{-1}. Additional short-distance data points are needed to determine if the function accurately describes the data set.

in water (A and k_0 converging), possibly owing to decreased chain stiffness in Gdn [64, 68]. Including Ru66 (pH 7) data in a fit to the extended equation with constrained A and y parameters gives $k_0 = 2.3 \times 10^7$ s^{-1}, or an upper time constant of 43 ns for close contacts (Figure 4.11).

4.6 Conclusions

The upper limit for the kinetics of the folding reaction will be governed by the rate at which intrachain contacts are formed. We measure a limiting time constant of 0.1 μs in 6 M Gdn, extrapolated to an estimated 30 ns in water. We observe a power-law dependence of contact formation rate constants on n, the length of the polypeptide chain between the contacts. Adherence of our data to a slope of -1.5, consistent with theoretical models for ideal polypeptides, demonstrates that cytochrome cb_{562} is minimally frustrated (loops 3–4). Contacts that stabilize the native structure are thus strongly favored during folding. Early formation of these native contacts can aid the protein in forming additional contacts, as we have observed, by effectively decreasing n. These events are expected to increase in occurrence as the folding reaction proceeds, in accord with a funneled energy landscape [8, 11], to the native state.

4.7 Acknowledgments

This work was carried out in collaboration with Dong Woo Shin, a Caltech student. We would like to thank the following former and current group members for the synthesis of the Ru label: Lionel Cheruzel, Jeffrey Warren, Maraia Ener, and Katja Luxem.

Chapter 5

Resolving the Fast Folding of Cytochrome cb_{562} with Microfluidic Mixing

5.1 Introduction

Cytochrome cb_{562} belongs to a unique family of cytochromes with similar topologies but divergent folding kinetics. We would like to fully characterize the folding mechanisms of cytochromes cb_{562} and c' and isolate the factors causing the differences. In this chapter, we resolve the folding kinetics of ferricytochrome cb_{562} by trFRET and continuous flow mixing.

Previous folding studies of ferrocytochrome cb_{562} were initiated by electron injection and monitored by transient absorption spectroscopy [36]. Met7 heme coordination occurs within 10 μs, and the absorption spectrum resembles the folded state after 100 μs. This experimental result suggests that persistent heme ligation aids in the folding process. Folding of the ferricytochrome is not expected to be as rapid as the ferrocytochrome since Met7 is more strongly coordinated to

the heme in the reduced protein.

Folding of ferricytochrome cb_{562} was investigated by stopped-flow mixing [37]. When monitored by Trp59 fluorescence, more than half of the signal is quenched during the 5 ms mixing dead time. The remaining amplitude decays with a denaturant-dependent rate constant that extrapolates to 4.2 x 10^2 s^{-1} at 0 M Gdn. Absorption-monitored kinetics are biphasic and include the fast phase observed with fluorescence and an additional denaturant-independent slow phase ($k_{obs} \approx 5$ s^{-1}). This slow phase likely involves optimization of the native heme environment, including the Met7 ligation.

Microfluidic continuous flow mixing (150-μs dead time) allows for the resolution of the early folding of cytochrome cb_{562} [29]. Trp59(helix3)–heme distance distributions were probed throughout folding by trFRET. Helices 3–4 fold with a rate constant of 240–260 s^{-1}. Interestingly, cytochrome c' also exhibited apparent two-state folding when Trp32–heme and Trp72–heme distances were probed, but differing rates [28]. This probe dependence of the rate constants are indicative of a sequential folding mechanism where the region around the helix 1/2 loop forms native contacts more rapidly (k = 170 s^{-1}) than the helix 2/3 loop (k= 15 s^{-1}).

In this chapter, we investigate the fast folding of cytochrome cb_{562} with probes that report on helix 1–heme and helix 3–heme folding. Dansyl (Dns) fluorophores are attached to K19C or D66C in single-cysteine variants (Figure 5.1). By comparing the rate constants for Dns19- and Dns66-cytochrome cb_{562}, we learn whether the protein folds with a concerted or sequential mechanism.

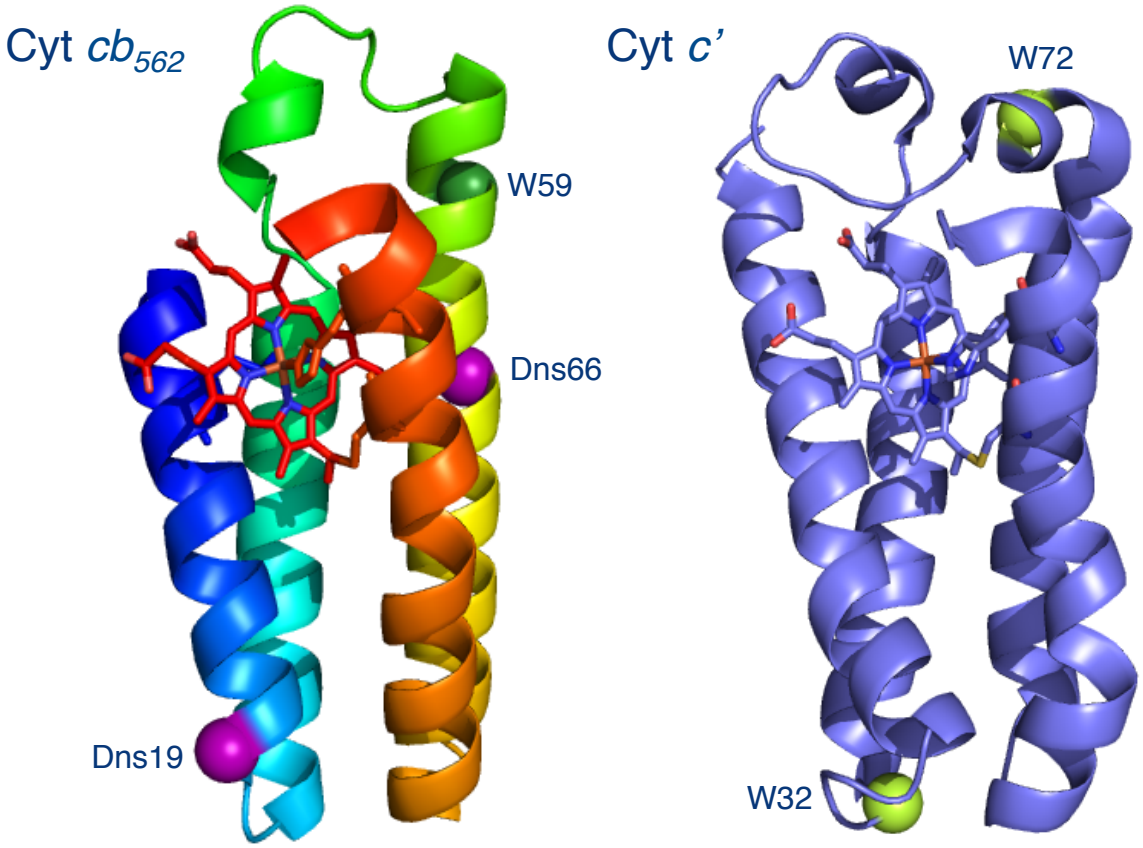

Figure 5.1: Cytochrome cb_{562} with helix I to helix IV/heme colored from blue to red (PDB 2BC5) and cytochrome *c'* (PDB 1MQV). Dansyl-labeling positions and tryptophan residues are highlighted as purple and green spheres, respectively. Folding is measured as a function of fluorophore to heme distance.

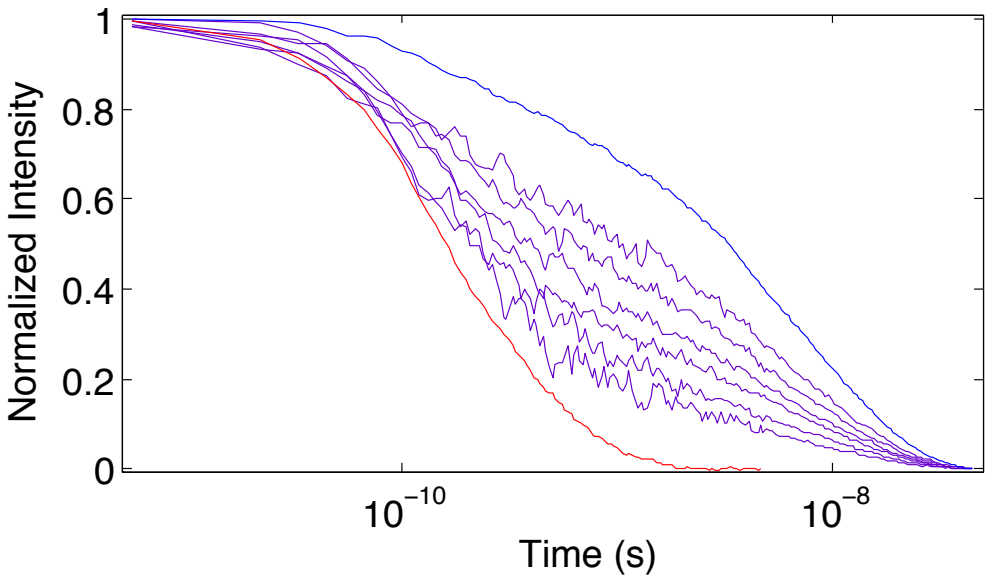

Figure 5.2: Dns66 fluorescence decays: unfolded (blue), during the folding reaction (purple), and folded (red). Dns66-cytochrome cb_{562} folding was triggered by Gdn jump from 6 to 1 M in a continuous flow mixer (pH 4.1, ambient temperature).

5.2 Dansyl-D66C: Folding of Helices 3-4

We triggered refolding using a continuous flow mixer to dilute denatured protein ([Gdn]= 6 M) with folding buffer ([Gdn]= 0 M) on a sub-millisecond timescale ($[\text{Gdn}]_{\text{final}}$= 1 M) and monitored the reaction progress using trFRET. Measurements were performed at pH 4 to prevent heme misligation by His63. The Dns66 fluorescence decay rate increases as the protein folds, indicating more efficient energy transfer to the heme (Figure 5.2). P(r_{DA}) distributions for Dns66 were extracted from fitting the fluorescence decay curves (Figure 5.3). Populations of intermediate and extended states decrease over time as additional compact structures are formed.

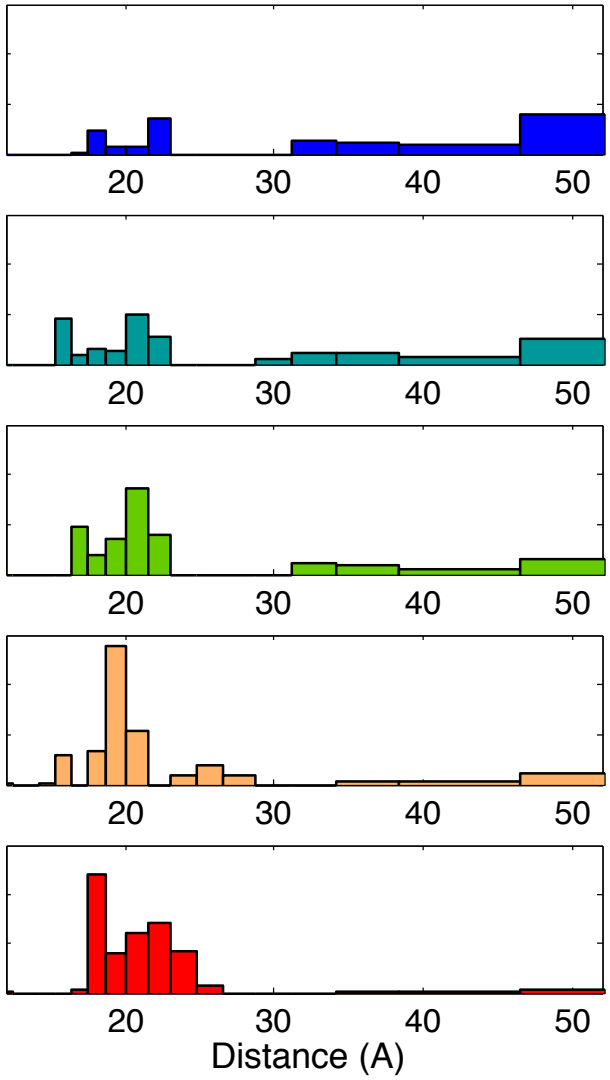

Figure 5.3: Conformational changes of cytochrome cb_{562} during folding, probed by Dns66-heme distances: unfolded (blue), after 1.5 ms (teal), after 4 ms (green), after 6 ms (orange), and folded (red).

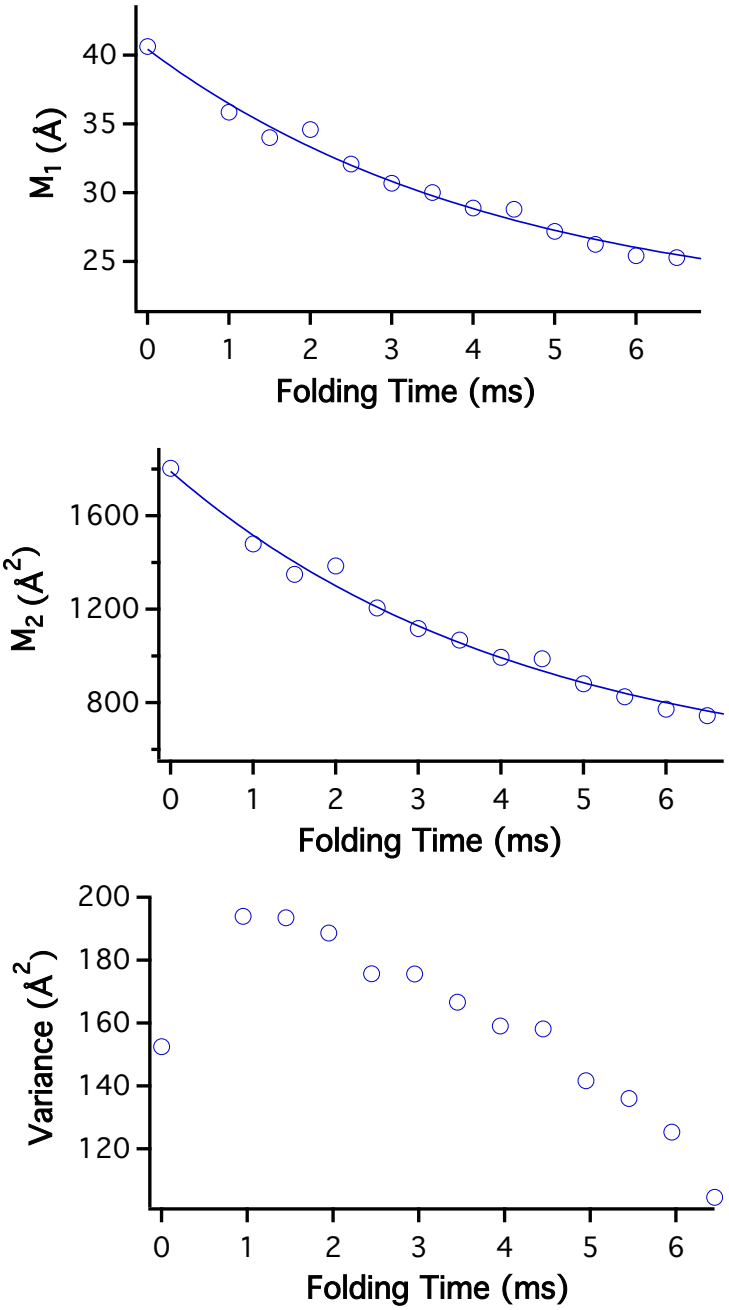

Figure 5.4: Analysis of the P(r_{DA}) fits of the folding kinetics of Dns66-cytochrome cb_{562}. (A) Mean distance between Dns66 and the heme (M_1) as a function of folding time. The solid line is a monoexponential fit with a rate constant of 230 s^{-1}. The infinite-time value of M_1 for the folded protein (22 Å) was included in the fit. (B) Time course of the second moment (M_2). The solid line is a monoexponential fit with a rate constant of 230 s^{-1}. The infinite-time value of M_2 for the folded protein (500 $Å^2$) was included in the fit. (C) Time course of the variance. The variance for the folded protein is 33 $Å^2$. For a two-state folding process, the variance should be biphasic with a growth rate constant equal to $2k_{obs}$ and a decay equal to k_{obs} [29].

The kinetics of P(r_{DA}) evolution were evaluated by moment analysis (see Methods 2.5). For a two-state process, the time course of the first (M_1) and second (M_2) moments will be exponential, with a rate constant corresponding to k_{obs} for the reaction [29]. We measure a k_{obs} of 230 s^{-1} (Figure 5.4A/B). The folded M_1 and M_2 were included in the fits; the trendlines agree well with the folding data and extrapolate to the moments of the native protein. The time course of the variance (V) in a two-state transition is biphasic, with a growth rate constant equal to $2k_{obs}$, and a decay equal to k_{obs} [29]; two phases are observed, consistent with a two-state process (Figure 5.4C). These results indicate that helix 3 forms native contacts with helix 4 by a two-state transition from the unfolded state ensemble to the native state with a rate constant of 230 s^{-1}.

5.3 Dansyl-K19C: Folding of Helices 1-4

Dns19 results parallel Dns66 findings. P(r_{DA}) distributions for the refolding of Dns19 (Figure 5.6) were extracted from fitting the fluorescence decay curves (Figure 5.5). An increase in energy transfer efficiency is observed throughout folding, consistent with the compaction of protein conformations. The unfolded state ensemble of Dns19 consists almost exclusively of extended structures, which gradually contract to intermediate then compact states during folding.

The kinetics of P(r_{DA}) evolution were evaluated by moment analysis. We fit the time courses of M_1 and M_2 to k_{obs} of 220 and 230 s^{-1}, respectively (Figure 5.7A/B). The folded M_1 and M_2 were included in the fits; the fits are consistent with the

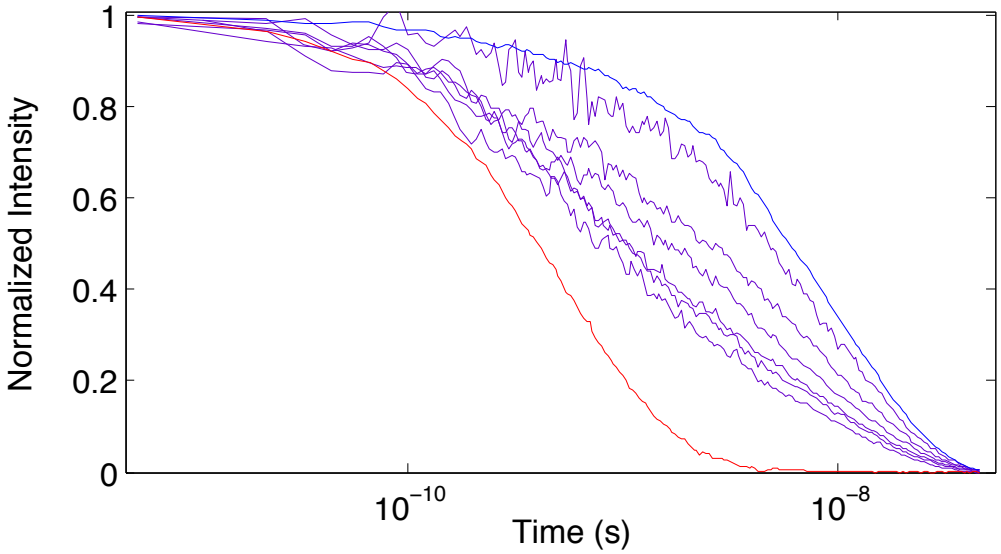

Figure 5.5: Dns19 fluorescence decays: unfolded (blue), during the folding reaction (purple), and folded (red). Dns19-cytochrome cb_{562} folding was triggered by Gdn jump from 6 to 1 M in a continuous flow mixer (pH 4.1, ambient temperature).

moments of the native protein and partially folded ensembles. Two phases are observed for V, as predicted for a two-state process (Figure 5.7C). These results indicate that the protein structure between the helix 1/2 loop and the heme folds by a two-state transition with a rate constant of 220–230 s^{-1}.

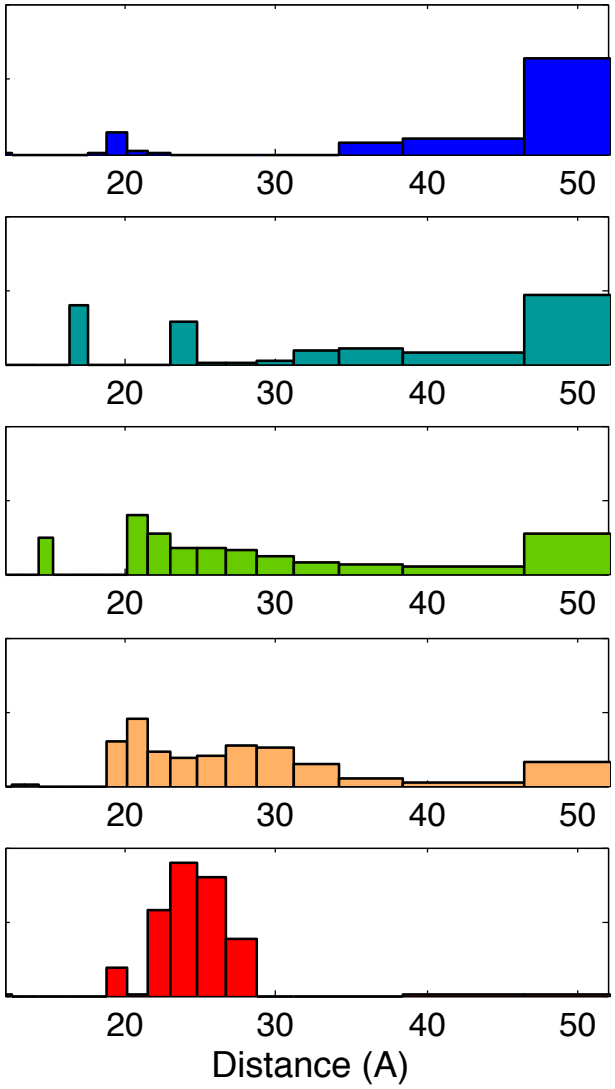

Figure 5.6: Conformational changes of cytochrome cb_{562} during folding, probed by Dns19-heme distances: unfolded (blue), after 1.3 ms (teal), after 3.3 ms (green), after 5.3 ms (orange), and folded (red).

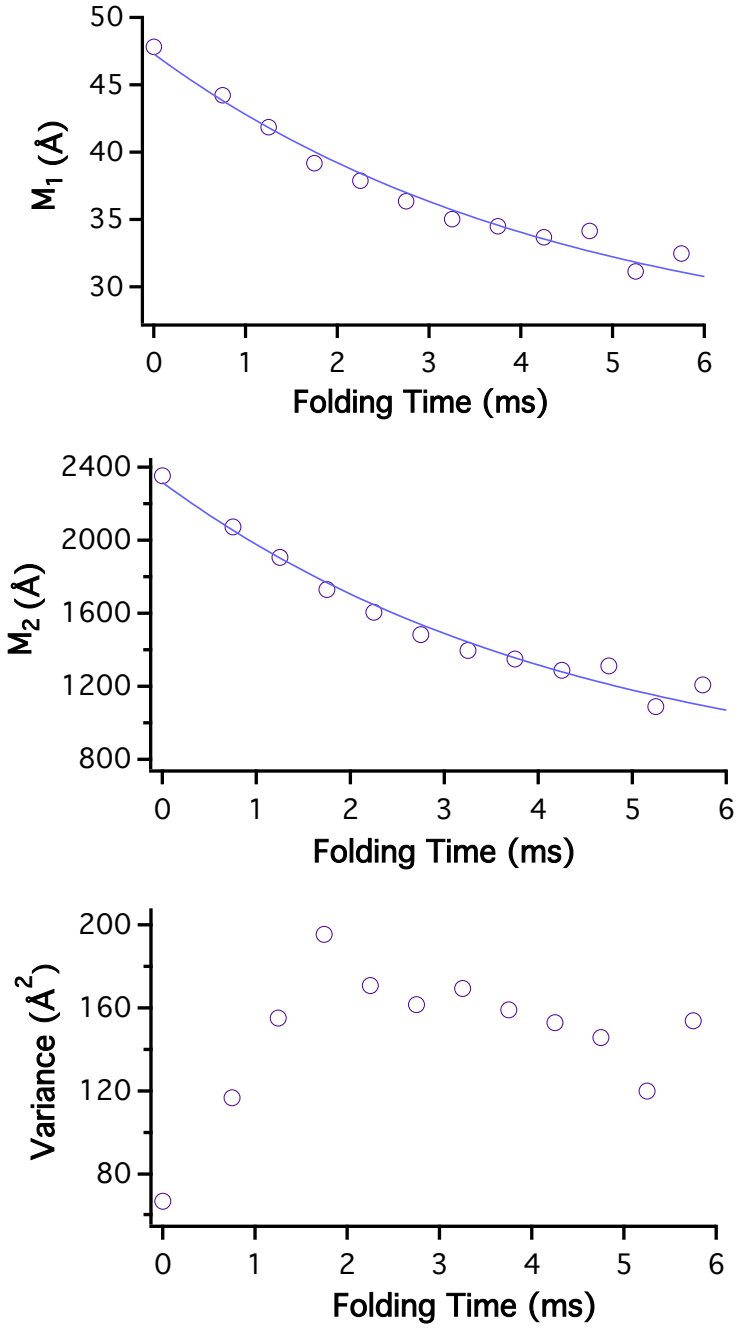

Figure 5.7: Analysis of the P(r_{DA}) fits of the folding kinetics of Dns19-cytochrome cb_{562}. (A) Mean distance between Dns19 and the heme (M_1) as a function of folding time. The solid line is a monoexponential fit with a rate constant of 220 s^{-1}. The infinite-time value of M_1 for the folded protein (25 Å) was included in the fit. (B) Time course of the second moment (M_2). The solid line is a monoexponential fit with a rate constant of 230 s^{-1}. The infinite-time value of M_2 for the folded protein (620 $Å^2$) was included in the fit. (C) Time course of the variance. The variance for the folded protein is 11 $Å^2$. For a two-state folding process, the variance should be biphasic with a growth rate constant equal to $2k_{obs}$ and a decay equal to k_{obs} [29].

Cyt cb_{562}: A D [L E D N M E T L N D N L K V I E K] A D N 22
Cyt c': A T D [V I A Q R K A I L K Q M G E A T K P I A A M L K] G E A K F D 33

[A A Q V K D A L T K M R A A A L D A] Q K A T P P K L E D K S P D S 55
[Q A V V Q K S L A A I A D D S K K] L P A L F P A D S K T G G D T A A L P 69

[P E M W D F R H G F D I L V G Q I D D A L K L A N] E G K 83
[K I W E D K A K F D D L F A K L A A A A T A A Q G T] I K D 98

[V K E A Q A A A E Q L K T T C N A C H Q K Y] R 106
[E A S L K A N I G G V L G N C K S C H D D F] R A K K S 125

Figure 5.8: Sequences of cytochromes cb_{562} and c' with the helices separated by brackets. Hydrophobic residues are highlighted in purple. Dansyl-labeling sites are colored blue. Tryptophan residues (and Phe32 of the F32W/W72F cytochrome c' variant) are shown in green. Cysteine residues binding the heme and axial ligands are shown in red and orange, respectively.

5.4 Folding Mechanism of Cytochrome cb_{562}

The amino acid sequences of cytochromes cb_{562} and c' (Figure 5.8) give some insight into the divergence of folding kinetics. Cytochrome c' has a more hydrophobic sequence than cytochrome cb_{562} and shows evidence of a hydrophobic collapse, with one of the refolding populations more compact than the native state [28]. Hydrophobic collapse is not observed for cytochrome cb_{562}. Helix 1 folds on a similar timescale in both proteins (4 ms), but helices 2–3 fold significantly slower in cytochrome c' (70 ms) than in cytochrome cb_{562} (4 ms). Collapse of hydrophobic patches in these regions could slow formation of helices and native intermolecular contacts [72, 73].

Frustration calculations [74] for native contacts in these proteins are consistent with our kinetics data. Both cytochromes cb_{562} and c' show a large degree of frustration between helix 1 and helix 2 (Figure 5.9). However, cytochrome c' has

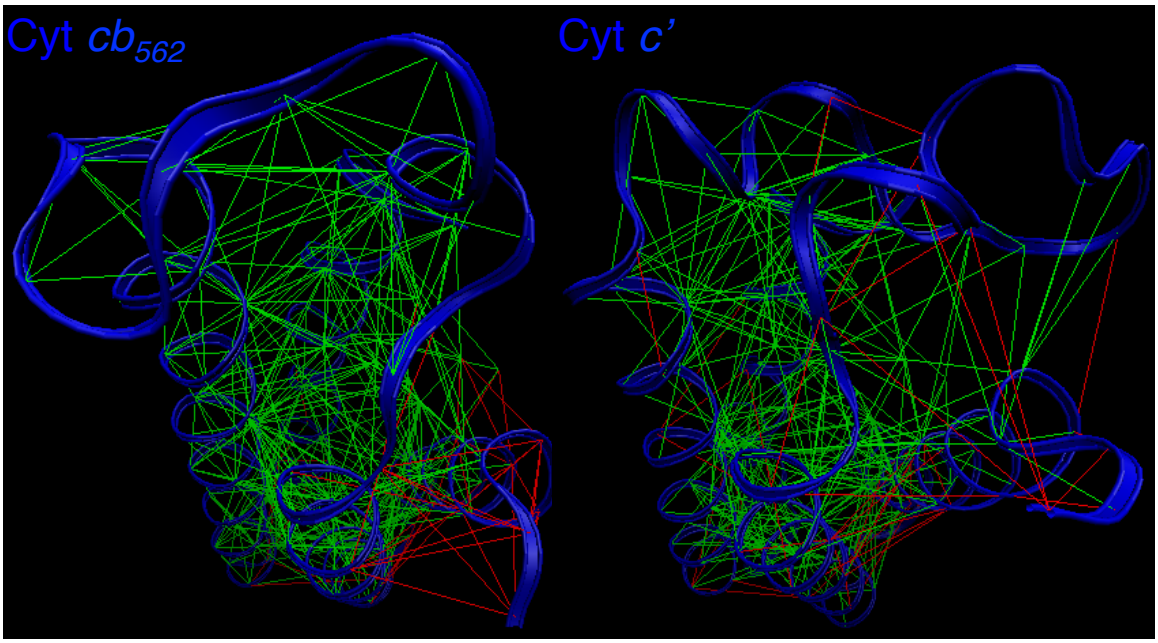

Figure 5.9: Top-down view of cytochromes cb_{562} (PDB 2BC5) and c' (PDB 1MQV) with the loop region in the foreground and the N-terminus in the bottom right corner. Frustrated contacts and minimally frustrated contacts are highlighted in red and green, respectively. The methodology is described in reference [74].

additional frustrated contacts within the helix 2/3 loop region and between the loop and the N-terminus. This indicates that these contacts are not energetically optimized, and nonnative contacts may be more favorable, which would slow folding. Cytochrome cb_{562} folding in this region is expected to be faster than for cytochrome c'.

Folding kinetics of cytochrome cb_{562} variants Dns19, Dns66, and Trp59 are similar, indicative of a concerted folding mechanism. Our results provide strong support for cooperative, two-state folding. These fluorophores probe the majority of the protein, between the loop end of helix 1 and the heme, close to the C-terminus. The N-terminus could fold at a different rate than Dns19; however, the fast-phase rate constant observed by both heme absorption and Trp59 fluorescence

after the stopped flow mixing dead time suggests that the heme environment (including interactions with helix 1) forms during macromolecular assembly [37]. The modeled frustration in helix 1, thus, does not slow folding. An intermediate with helix 1 unfolded was observed after the rate-limiting transition state for engineered variants of, or "redesigned," apocytochrome b_{562} [69, 70, 71]. This hidden intermediate could also be involved in cytochrome cb_{562} folding, but may not be relevant to the heme-containing protein if the energy landscape is affected by the presence of the heme or the altered sequence.

Our experiments of helix 3 folding (Dns66) were carried out at pH 4 to avoid heme misligation in the unfolded protein (Chapters 3-4). Kinetics measurements for Trp59, nearby on helix 3, were carried out at pH 5 [29], where His63 is ligated to the heme in about 40% of the molecules. The presence of this contact does not significantly impact the folding rate; Trp59 probed folding is only marginally faster (k = 240–260 s^{-1}) than Dns66 folding (k = 230 s^{-1}). This robustness of the rate suggests that the contact does not divert molecules along a slower folding pathway. His63 is relatively close to the heme in the native protein (14 Å) and thus may not inhibit, and may even aid, native contact formation. The folding kinetics may also be unaffected if the conversion between the misligated and fully extended forms is fast. (The conversion time is unknown at 1 M Gdn, but greater than several microseconds at 6 M Gdn.) The energy landscape theory predicts malleable folding pathways in minimally frustrated proteins; therefore, subsequent folding events may not be affected by the early conformational

constraint.

5.5 Conclusions

Cytochrome cb_{562} exhibits cooperative two-state folding on the millisecond timescale, in contrast to cytochrome *c'*, which exhibits slower folding in the highly frustrated helix 2/3 loop. Folding kinetics studies reveal that the region between the helix 1/2 loop and the heme folds at the same rate as the region between helix 3 and the heme in a concerted folding event. The rate of helix 3 folding is unaffected by His63 ligation of the heme in the unfolded state. We have previously shown that helices 3–4 of cytochrome cb_{562} are minimally frustrated (Chapter 4); these additional studies suggest that the majority of the protein (helix 1/2 loop to the heme) is minimally frustrated.

5.6 Acknowledgments

The frustration calculations were performed by Patrick Weinkam while in Peter Wolynes' group at the University of California, San Diego.

Chapter 6

An Unfolding Intermediate of Cytochrome c_{552} Revealed with trFRET[1]

6.1 Introduction

Cytochrome c_{552} belongs to the class-I *c*-type cytochrome family, which includes yeast cytochrome *c*. Cytochromes in this family share the common structural feature where the N- and C-terminal helices cross each other (αI and αIV in Fig. 1A). However, cytochrome c_{552} has the distinct structural components of a β-sheet around the heme propionate and extra helices in the C-terminal region (Figure 6.1). The role of these extra structures has never been fully elucidated.

We are interested in the contribution of these distinct structural features to the folding kinetics and thermodynamic stability of this thermostable protein. We have probed the unfolding of cytochrome c_{552} under equilibrium unfolding conditions. The structural conformations are monitored using trFRET with

[1]Adapted from Yamada, S.*; Bouley Ford, N.*; Ford, W.; Keller, G.; Gray, H.; Winkler, J. "Snapshots of a Protein Folding Intermediate" *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 1606-1610 [51].

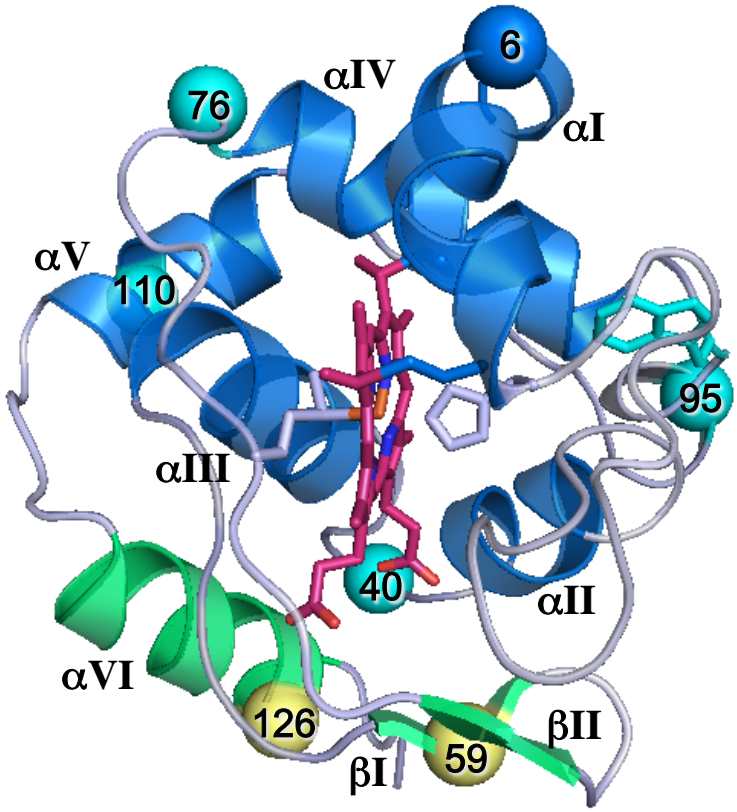

Figure 6.1: Crystal structure of WT cytochrome c_{552} at pH 5.44 (PDB code: 3VNW). Dns-labeled sites are shown as spheres: group I (cyan) and group II (yellow). Helix αVI and β-sheets I and II are distinct structural features of cytochrome c_{552} that exhibit group II folding behavior (green). For reference, Trp91 is also highlighted in cyan.

fluorescent probes distributed throughout the protein, including the region distinct to cytochrome c_{552} (Figure 6.1).

6.2 Stability of Labeled Mutants to Denaturation

Solvent-exposed residues that form no intramolecular bonds were chosen for trFRET studies to minimize structural perturbations upon labeling. Seven single-mutation variants of cytochrome c_{552} were prepared; in each mutant, a residue

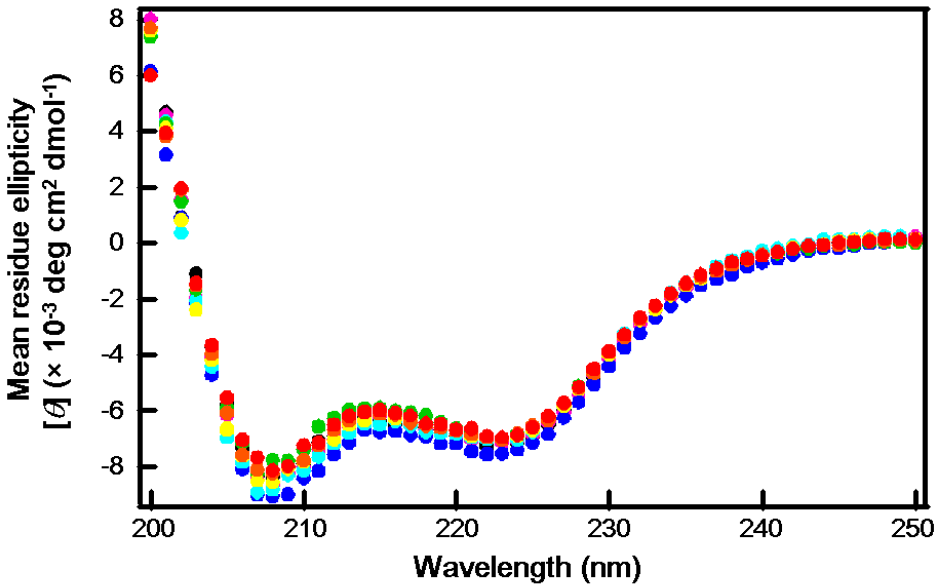

Figure 6.2: CD spectra of Dns cytochrome c_{552} variants. The spectra were measured with 20 μM protein in 10 mM sodium citrate buffer (pH 3.0) at 25°C: WT (black), Dns126 (magenta), Dns110 (blue), Dns95 (cyan), Dns76 (green), Dns59 (yellow), Dns40 (orange), and Dns6 (red).

(Lys6, Glu40, Glu59, Lys76, Lys95, Lys110 or Lys126) was mutated to cysteine and further modified by covalent Dns attachment.

Circular dichroism (CD) spectra of all Dns variants were virtually identical with that of WT (Figure 6.2), indicating that the native secondary structure was preserved in the Dns-labeled proteins. Denaturation curves of the Dns variants, obtained from Dns fluorescence, UV-visible heme absorption, and CD measurements, showed midpoints ($[\mathrm{Gdn}]_{1/2}$) close to 4 M guanidine hydrochloride (Gdn), consistent with that of WT (Table 6.1 and Figure 6.3). Standard denaturation parameters m and -ΔG_w [59], determined from these curves, also showed no substantial deviation from those of WT (Table 6.1). We conclude from these results that Dns labeling does not perturb the stability or secondary structure of any of the

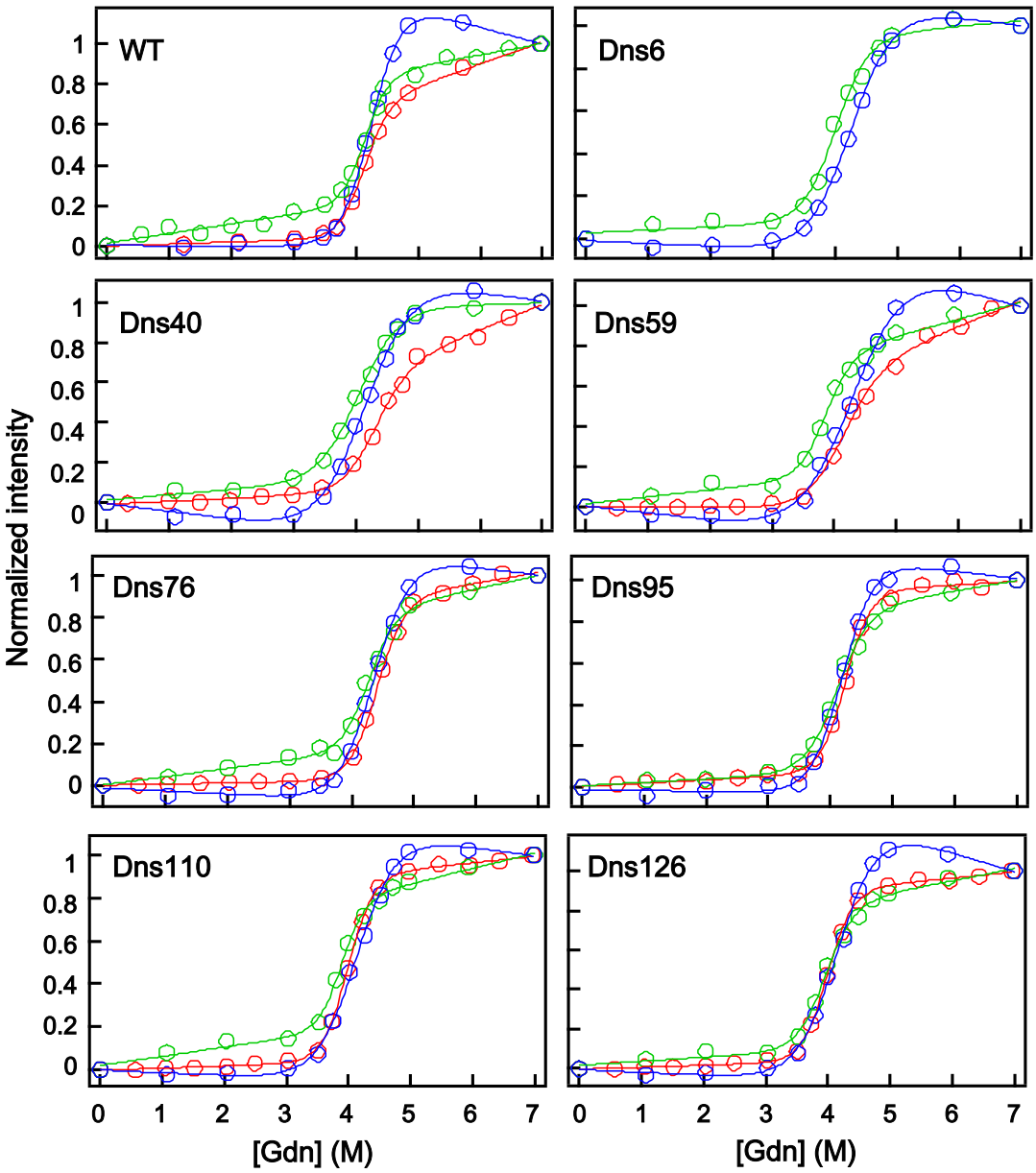

Figure 6.3: Denaturation curve of Dns cytochrome c_{552} variants. Normalized signal of steady state fluorescence (red), heme absorbance (blue), and circular dichroism (green) versus Gdn concentration. Fluorescence data for Dns6 is not provided because there was negligible difference between native and unfolded states.

variants.

6.3 Steady-State trFRET Measurements

We employed trFRET to acquire site-specific D-A (Dns-heme) distance distributions (P(r_{DA})) upon denaturation with Gdn as a method to detect unfolding intermediates. A Tikhonov regularization (TR) method was used to extract distance distributions (Figure 6.5) from the fluorescence decays shown in Figure 6.4. Measurements were performed at pH 3 to prevent misligation of the heme by histidine residues or the N-terminal amino group. In most cases, Dns fluorescence decays could not be fit to single exponential functions, as r_{DA} components exist with various probabilities, consistent with the presence of multiple conformations.

At 0 M Gdn, r_{DA} for all Dns variants obtained from TR fitting are consistent with the WT crystal structure (Figure 6.5 and Table 6.2). The most probable D-A distances extracted from the fluorescence decays (r_{mode}) are slightly longer (1-4 Å) than the Cγ-Fe distances in the crystal structure, likely owing to the length of linker between the Dns fluorophore and the Cγ atom.

Denaturation with Gdn slowed the fluorescence decay in all Dns variants (Figure 6.4), consistent with an increase in the average distance between D and A. Extended populations were observed in all but one of the Dns variants upon the addition of Gdn (Figure 6.5). Since Dns6 is located close to the heme and thus exhibited only a modest broadening of the D-A distribution when denatured (Figure 6.4), it will not be considered further.

Table 6.1: Denaturation parameters for Dns cytochrome c_{552} variants.

	Fluorescence*			Circular dichroism ‡			Absorption∓		
Variant	$[Gdn]_{1/2}$ (M)	m (kJ/mol/M)	$-\Delta G$ (kJ/mol)	$[Gdn]_{1/2}$ (M)	m (kJ/mol/M)	$-\Delta G$ (kJ/mol)	$[Gdn]_{1/2}$ (M)	m (kJ/mol/M)	$-\Delta G$ (kJ/mol)
WT	4.1±0.02†	11.9±0.8†	49±4†	4.1±0.03	13.2±1.6	54±7	4.3±0.02	11.9±0.8	49±4
Dns6	N/A	N/A	N/A	4.0±0.06	9.3±1.8	37±8	4.3±0.02	7.1±0.3	31±1
Dns40	4.3±0.05	9.1±1.2	39±6	4.1±0.04	7.6±0.8	31±4	4.2±0.05	6.4±0.6	27±3
Dns59	4.1±0.05	7.9±1.1	32±5	3.9±0.05	11.1±1.9	43±8	4.3±0.06	5.8±0.5	25±3
Dns76	4.4±0.02	11.0±0.7	48±3	4.3±0.05	11.0±2.0	47±9	4.4±0.03	8.6±0.7	38±3
Dns95	4.2±0.02	12.2±0.9	51±4	4.1±0.04	9.6±1.1	39±5	4.2±0.03	9.7±0.8	41±4
Dns110	4.2±0.02	12.8±1.5	51±6	3.9±0.03	13.2±2.0	51±8	4.1±0.03	7.7±0.6	32±3
Dns126	4.0±0.01	12.0±0.5	48±2	3.9±0.04	10.7±1.5	42±6	4.2±0.03	7.7±0.6	32±3

Measured at 25°C in pH 3.0 citrate buffer (10 mM for circular dichroism and 100 mM for fluorescence and heme absorption).
∗*Steady-state measurements of Dns fluorescence intensities. Excitation and emission wavelength are 355 and 513 nm, respectively.
†Data of WT were taken from Trp91 fluorescence. Excitation and emission wavelength are 290 and 360 nm, respectively.
‡Ellipticity measurements at 222 nm.
∓Heme absorbance measurements at 391 nm.

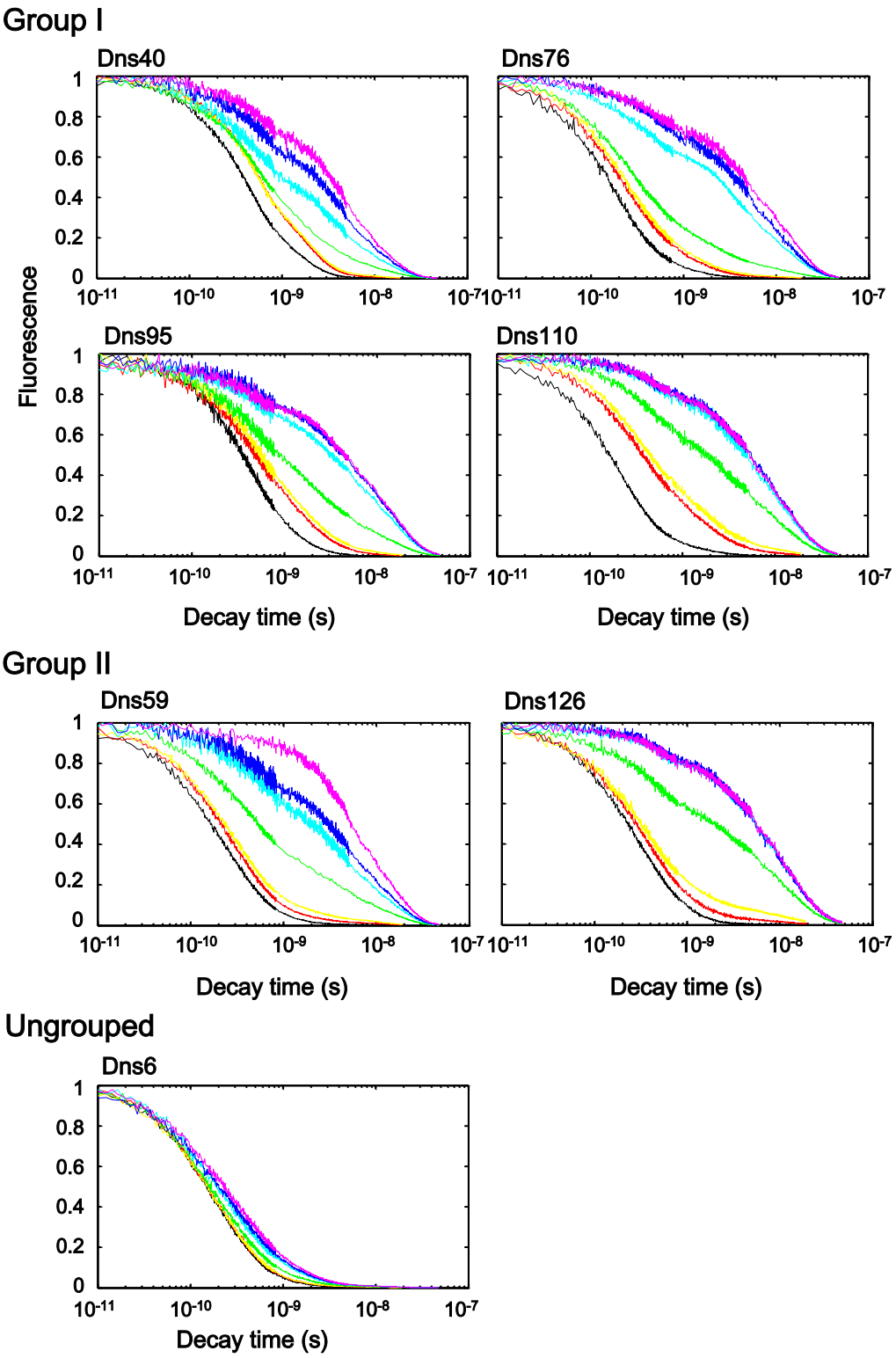

Figure 6.4: Fluorescence decay curves of Dns cytochrome c_{552} variants upon Gdn-denaturation: Gdn 0 M (black), 3 M (red), 3.5 M (yellow), 4 M (green), 4.5 M (cyan), 5 M (blue), and 7 M (magenta).

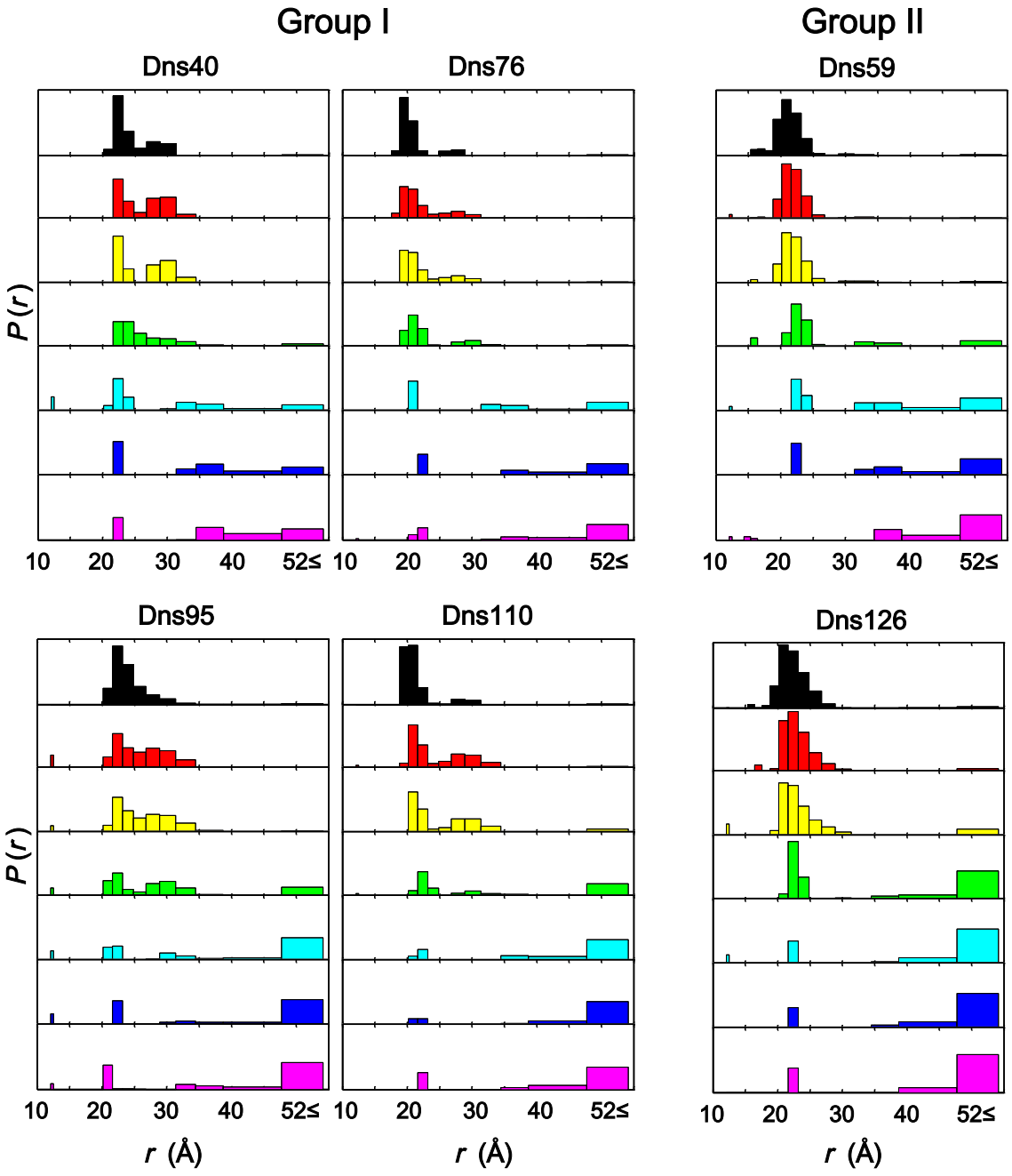

Figure 6.5: Gdn-induced changes in the Dns-heme distance distributions [(P(r_{DA}))] for Dns cytochrome c_{552} variants at pH 3.0, extracted from TR fitting of fluorescence decay curves (Figure 6.4): Gdn 0 M (black), 3 M (red), 3.5 M (yellow), 4 M (green), 4.5 M (cyan), 5 M (blue), and 7 M (magenta). The area of each bar reflects the probability amplitude over the corresponding distance range. Group I and group II variants display three-state and two-state folding transitions, respectively.

Table 6.2: Native-state donor-acceptor distances for dansyl-labeled cytochrome c_{552} variants.

Variant	Distance C_γ-Fe (Å)*	r_{mode} (Å)†	r_{mean} (Å)†
6	17.5	21	22
40	22.7	22	25
59	16.3	21	22
76	18.9	19	22
95	21.7	22	25
110	21.6	21	22
126	16.7	21	23

∗Measured from the crystal structure (PDB 3VNW).
†Extracted from the Dns-heme distance distributions.

We defined three categories of structural conformations from the distance distributions: native/compact (C: $18 \leq r_{DA} \leq 22$ Å); intermediate (I: $25 \leq r_{DA} \leq 35$ Å); and extended (E: $r_{DA} \geq 35$ Å), and we plotted their relative populations as functions of [Gdn] (Figure 6.6). Population of E conformations in all variants occurs at Gdn concentrations greater than 4 M, which is consistent with the denaturation curves determined using absorption, CD, and steady-state fluorescence spectroscopic methods (Figure 6.3). Interestingly, the Dns variants can be divided into two groups that differ in how C, I, and E populations change with Gdn concentration; Dns40, 76, 95, and 110 are in group I, and Dns59 and 126 are in group II (Figures 6.5 and 6.6).

C populations in group I transform into I components when denatured by 3 to 4 M Gdn, which then disappear at higher concentrations ($\geq$ 5 M, Figure 6.6). The findings for group I variants are consistent with a three-state transition among the C, I, and E populations. The appearance and disappearance of I structures indicate that they arise from an unfolding intermediate state and are not simply

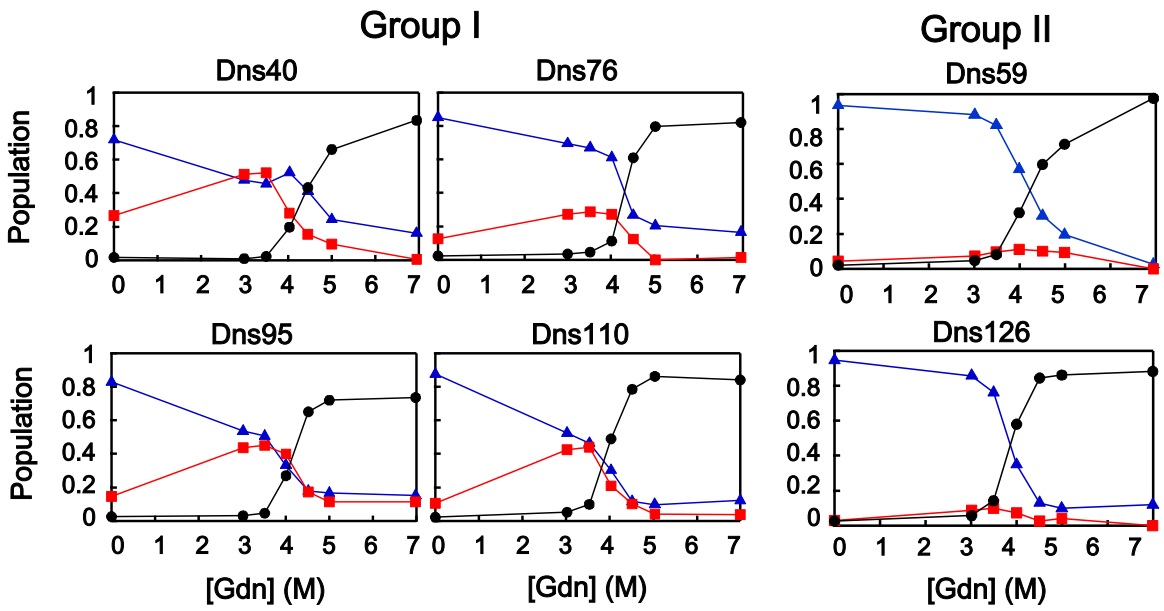

Figure 6.6: Population changes of structural conformations as a function of the Gdn concentration: native/compact (N), blue triangle; intermediate (I), red square; and extended (E), black circle. Group I and group II variants display three-state and two-state folding transitions, respectively.

constituents of the denatured protein ensemble. The I state has a folding stability similar to that of the native structure as I appears at relatively low [Gdn] in the presence of C (Figure 6.5). Absorption spectroscopy indicates that the heme environment in the intermediate is structurally similar to the native state; the heme Soret peak maxima are comparable at 0 and 3 M Gdn, indicating that the heme is low-spin with an intact, native Fe-S(Met) ligation (Figure 6.7). It is possible that I corresponds to a local minimum on the native protein folding energy landscape. In the absence of denaturant, its population is too small to be detected, but the introduction of a low concentration of denaturant shifts the equilibrium to populate this extended conformation rather than the compact native structure.

With fluorescent probes at Dns59 and Dns126, we found no evidence for

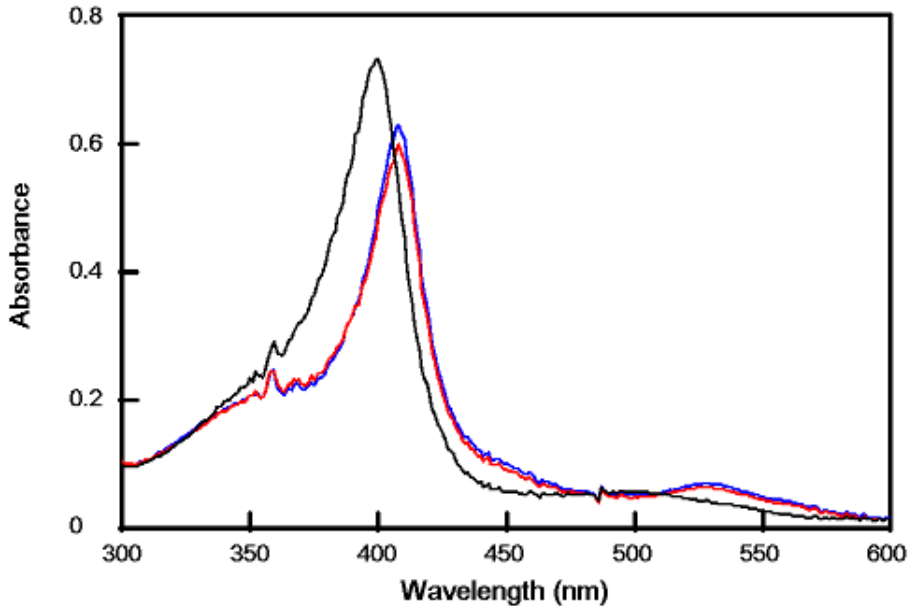

Figure 6.7: Absorption spectra of Dns110-cytochrome c_{552} in Gdn: Gdn 0 M (blue), 3 M (red), and 7 M (black). The spectra of 4.9 μM protein were collected in 10 mM sodium citrate buffer (pH 3) at 25°C.

the existence of I components during Gdn denaturation; these two derivatives form group II (Figure 6.6). Extended structures appear in both groups around 4 M. If we exclude the possibility that mutation and labeling at residues 59 and 126 have altered the unfolding pathway, then we can conclude that the unfolding intermediate formed at low Gdn concentrations involves little structural rearrangement from the native structure in the vicinity of Dns59 and Dns129 with slight expansion in the rest of the protein (Figure 6.8).

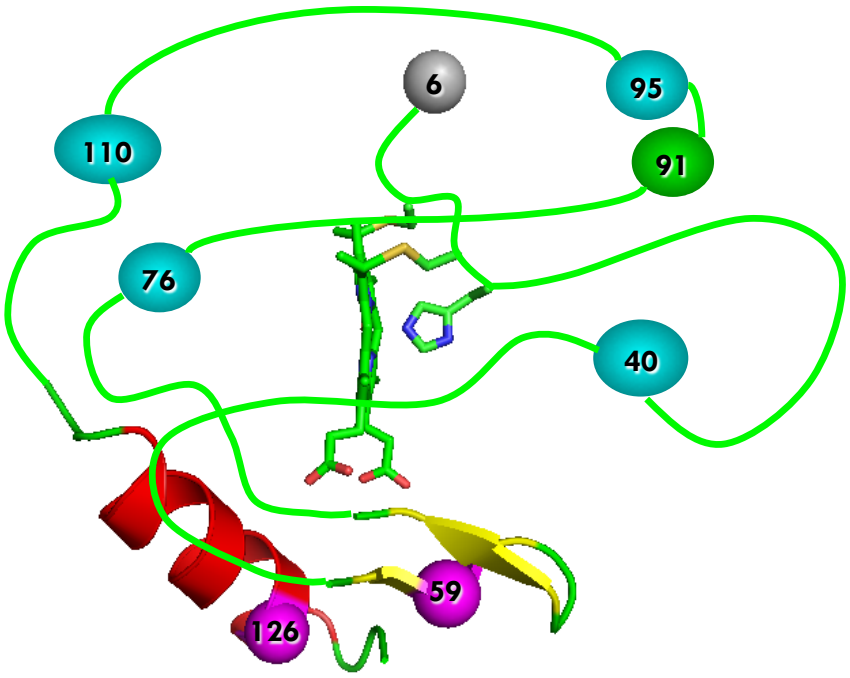

Figure 6.8: Schematic of proposed equilibrium unfolding intermediate with group II variants at native distances and group I variants at intermediate distances from the heme.

6.4 Cytochrome c_{552}'s Distinct Structural Features

Interestingly, the two group II residues are located in the distinct protein structural region that differentiates cytochrome c_{552} from other class-I *c*-type cytochromes (position 59 on β-sheet βI and 126 on the extra helix αVI). The extra helix αVI and the β-sheet βI of cytochrome c_{552} form a network of interactions (via residues 59, 125, and 131) linking them to the heme propionate (Figure 6.9). The WT crystal structure reveals that the amide nitrogen and carbonyl oxygen of Glu59 can hydrogen bond with the Lys131 carbonyl oxygen and the Arg125 side chain, respectively. Additionally, the Arg125 side chain forms salt bridges with the heme propionate and the carbonyl oxygen of Lys131. It is likely that these interactions stabilize this region of the polypeptide chain near the heme, making it more resistant to denaturation than the parts of the protein containing group I residues. By way of contrast, the other distinct helix in cytochrome c_{552} (αV) does not exhibit enhanced stability, insofar as the Dns110 label, located in αV, displays a three-state folding transition (Figures 6.5 and 6.6).

In the original report of the *T. thermophilus* cytochrome c_{552} x-ray crystal structure, it was proposed that extra structural elements in the protein were responsible for its stability at high temperatures [75]. The inherent stability of the cytochrome c_{552} fold is reflected by the finding that the apoprotein retains the same secondary structure as the holoprotein [76]. In marked contrast, the apoprotein of mitochondrial cytochrome *c* is unstructured; in this case, the heme, which acts as a hydrophobic nucleation site, is needed to stabilize the native fold [77]. Our

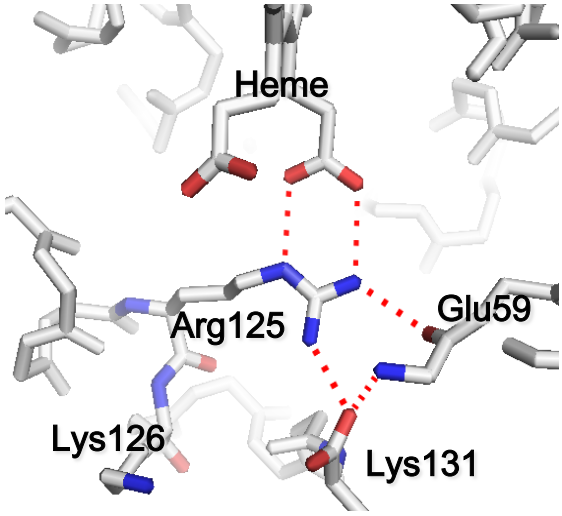

Figure 6.9: Crystal structure of WT cytochrome c_{552} obtained at pH 5.44 (PDB code: 3VNW). Red dashed lines indicate electrostatic interactions in the region of distinctive structures of cytochrome c_{552}.

cytochrome c_{552} unfolding data confirm that the extra C-terminal helices and β-sheet enhance the stability of the native structure by stabilizing the heme region.

Analogous features in yeast cytochrome c behave similarly under equilibrium unfolding conditions. In our work on yeast cytochrome c, I components were observed at moderate-to-high Gdn concentrations (2.7–5.9 M Gdn) for Dns39 and 50 near the heme propionate site and for Dns66 near the Met80 axial ligand site [21], indicating that compact structure is preserved in the heme region when unfolded. Our work on cytochrome c_{552} reveals that heme region variants (Dns59 and 126) do not develop I or E populations at mild Gdn concentrations where intermediate structures are found in other regions of the protein (3–4 M Gdn, Figures 6.5 and 6.6). We conclude that β-sheet βI (Dns59) and helix αVI (Dns126) are substantially more resistant to denaturant-induced unfolding.

6.5 Conclusions

Our trFRET experiments have shed new light on region-specific conformational changes that occur in cytochrome c_{552} upon increasing denaturant concentration. Dansyl probes at seven different residues in the protein reveal the existence of a loosely packed unfolding intermediate structure. We have also revealed that two sites associated with the distinct structural features of this class-I cytochrome are more resistant to unfolding than the rest of the protein.

6.6 Acknowledgments

This work was carried out in collaboration with Seiji Yamada, a visiting scientist from Sony Corporation [51].

Chapter 7

Characterization of Cytochrome c_{552}'s Folding Intermediate[1]

7.1 Introduction

Folding intermediates can lead to misfolding; however, certain intermediates can actually help the protein find its native structure by avoiding traps in the folding energy landscape [8, 9, 78, 21]. These states are referred to as on-pathway intermediates [79].

An on-pathway folding intermediate of cytochrome c_{552} was identified by Brunori and coworkers [49]. The native Trp91 fluorescence quenching during refolding was biphasic. Rate constants for refolding and unfolding were measured by stopped-flow mixing and double-mixing experiments at different Gdn concentrations. Rate constants for the fast unfolded↔intermediate and slow intermediate↔native transitions in the absence of denaturant were extracted from chevron plots: k_{UI} = 1100 s^{-1}, k_{IU} = 0.2 s^{-1}, k_{IN} = 2 s^{-1}, and k_{NI} = 4 x 10^{-6} s^{-1}, where U,

[1]Adapted from Yamada, S.*; Bouley Ford, N.*; Ford, W.; Keller, G.; Gray, H.; Winkler, J. "Snapshots of a Protein Folding Intermediate" *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 1606-1610 [51].

I, and N are the unfolded, intermediate, and native states, respectively.

We have further characterized the refolding kinetics for the fast formation of this intermediate by microfluidic continuous flow mixing and trFRET. A strength of trFRET is that it allows us to characterize heterogeneity in ensembles of protein conformations [80].

7.2 Folding Kinetics

We triggered refolding using a continuous flow mixer to dilute denatured protein ([Gdn]= 6 M) with folding buffer ([Gdn]= 0 M) on a sub-millisecond timescale ($[\text{Gdn}]_{\text{final}}$= 1 M) and monitored the reaction progress using trFRET. To monitor the kinetics of cytochrome c_{552} refolding, we selected the Dns110 derivative (located at helix αV) as representative of the three-state transition folding variants from group I. We also measured the refolding of Dns59 (data not shown); denaturation of this variant resulted in partial aggregation.

The Dns110 fluorescence decay rate increases as the protein folds, indicating more efficient energy transfer to the heme (Figure 7.1A). Very little compaction of the protein occurs on timescales shorter than 200 μs. The distance distributions ($P(r_{DA})$, Figure 7.1B) reveal intermediate ($25 \leq r_{DA} \leq 35$ Å) and compact ($18 \leq r_{DA} \leq 22$ Å) components developing in the first millisecond of folding. Compact populations gradually increase with time, but some extended ($r_{DA} \geq 35$ Å) and intermediate structures still remain at 3.8 ms, the longest measurable reaction time.

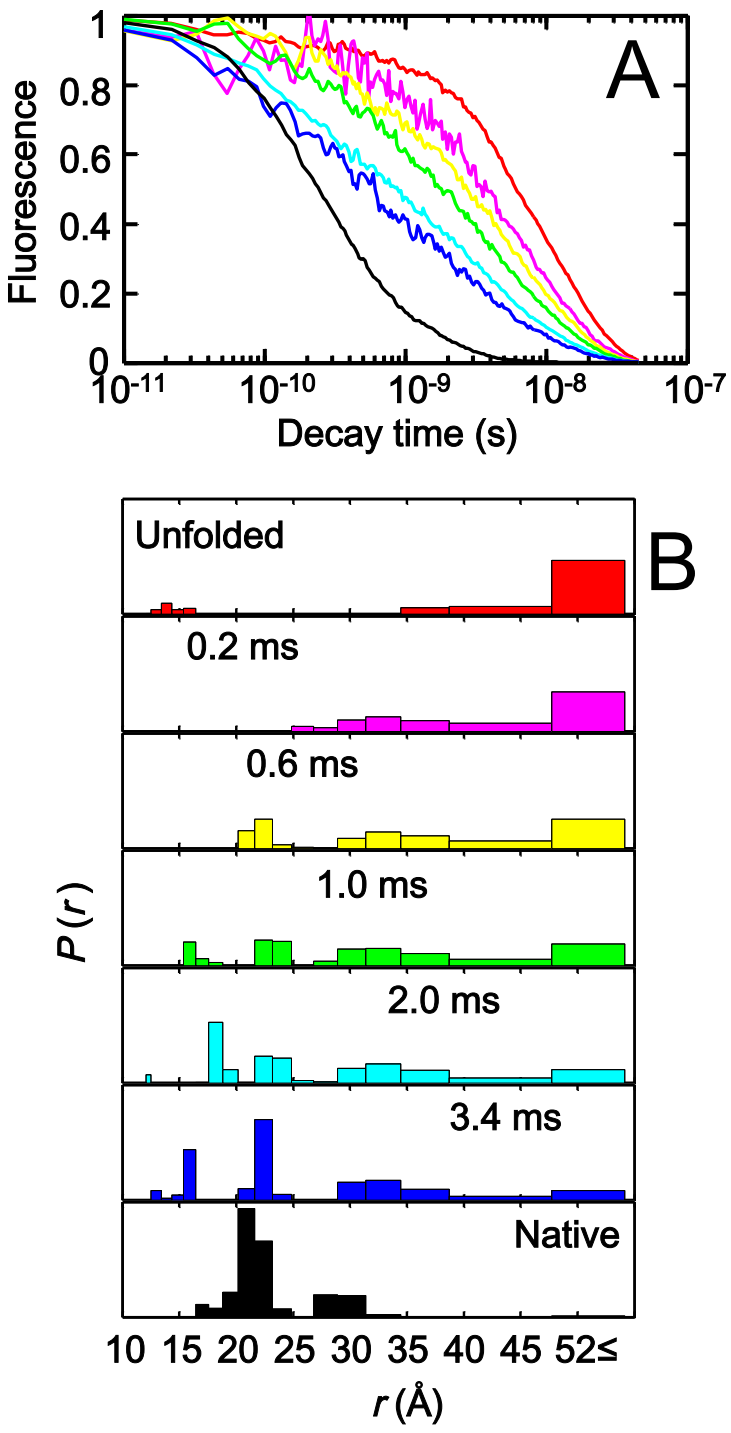

Figure 7.1: Dns110-cytochrome c_{552} folding triggered by Gdn jump from 6 to 1 M in a continuous flow mixer (pH 3.0, ambient temperature). (A) Fluorescence decay curves. For clarity, only 7 of the 23 observed decays are displayed: unfolded in 6 M Gdn (red), 0.2 (magenta), 0.6 (yellow), 1.0 (green), 2.0 (cyan), 3.4 ms (blue), and folded in 1 M Gdn (>30 min after initiation of folding reaction, black). (B) P(r_{DA}) distributions for refolding extracted from fitting the fluorescence decay curves.

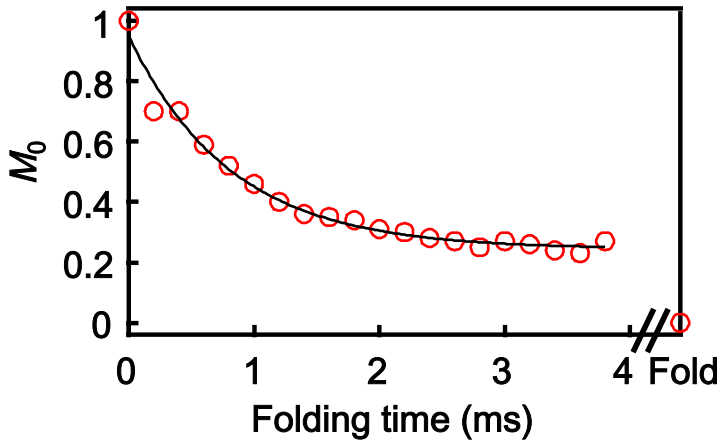

Figure 7.2: Folding kinetics of Dns110-cytochrome c_{552} triggered by Gdn jump from 6 to 1 M in a continuous flow mixer (pH 3.0, ambient temperature). Integrated Dns-cytochrome c_{552} fluorescence intensity (M_0) is shown as a function of folding time. The solid line is a monoexponential fit with a rate constant of 1230 s^{-1}.

The kinetics of P(r_{DA}) evolution were evaluated by moment analysis (see Methods 2.5). The time course of the normalized integrated fluorescence intensity (M_0) can be fit to a monoexponential function (k_{obs} = 1230 s^{-1}, Figure 7.2). For a two-state process, the time course of the first (M_1) and second (M_2) moments will be exponential, with a rate constant corresponding to k_{obs} for the reaction [29]; we find k_{obs} of 1060 and 1140 s^{-1} from the time courses of M_1 and M_2, respectively (Figure 7.3A/B). The time course of the variance (V) in a two-state transition is biphasic, with a growth rate constant equal to $2k_{obs}$, and a decay equal to k_{obs} [29]; we find that V agrees with this model (k_{obs} = 1100 s^{-1}, Figure 7.3C). These results indicate that Dns110 exhibits a two-state transition with a rate constant of 1100 s^{-1} from the unfolded conformations.

Interestingly, the millisecond folding process does not appear to produce the

native state, as roughly one fourth (24%) of the normalized integrated fluorescence intensity persists at the end of this kinetics phase (Figure 7.2). Since the extrapolated endpoints to the fits of the moments (M_1 = 31.5 Å, M_2 = 1200 Å^2, and V = 200 Å^2) are not in accord with values for the native protein (M_1 = 23 Å, M_2 = 560 Å^2, and V = 31 Å^2, Figure 7.3), the millisecond process must correspond to formation of an intermediate folding ensemble. Notably, this ensemble has a greater population of extended components (Figure 7.1B, 3.4 ms) than that observed under equilibrium unfolding conditions. The early appearance of structures with 30 to 35 Å-Dns110-heme distances suggests partial collapse involving the heme and helix αV, but the scarcity of 20-Å structures indicates little native-state formation.

This refolding intermediate is distinct from the unfolding intermediate reported in the previous chapter. They two intermediates have different structural ensembles. The refolding intermediate forms rapidly and is partially comprised of extended structures, whereas the compact unfolding intermediate is structurally and energetically similar to the native state. The equilibrium unfolding intermediate, which contains native heme ligation, is not observed during refolding and thus occurs after the rate-determining step.

Brunori found evidence for an on-pathway folding intermediate in cytochrome c_{552} using native Trp91 fluorescence as a probe [49]. The difference between the previously reported rate constant for intermediate formation (k_{UI} = 400 s^{-1}, [Gdn]= 1 M) and ours can be rationalized by the differing pH conditions and temperatures employed (pH 2.1, 10°C versus pH 3.0, 18°C). The final step in

cytochrome c_{552} folding, the transition from intermediate to native structure, was reported to have a one-second time constant ([Gdn] = 1 M) [49]. The native protein contains five proline residues, all in the more stable trans configuration in the crystal structure. Brunori and coworkers observed a third refolding phase ($k \approx 10^{-2}$ s^{-1}) that, on the basis of interrupted unfolding experiments, they attributed to proline isomerization steps [49]. Hence, the early folding intermediate that we and Brunori have observed is not likely due to the presence of cis-Xaa-Pro traps.

Although the folding intermediate was proposed to be a single compact structure in which 60% of the exposed surface in the unfolded protein had been buried [49], our trFRET data suggest instead that it is a heterogeneous ensemble of compact and extended structures. Moment analysis of the Dns110 folding kinetics indicates that this ensemble has developed about 75% of native structure (Figure 7.2). For comparison, the burst phase ensemble of yeast cytochrome *c* contains a comparable amount of native character, 60-75%, depending on residue [78].

The remarkable difference between the two proteins is that, in yeast cytochrome *c*, this structure develops in less than 150 μs [25], whereas comparable structure development in cytochrome c_{552} is almost ten times slower. The sluggish refolding of cytochrome c_{552} carries over into the final phase as well, as a full second is required for formation of the native state; in yeast cytochrome *c*, the folded state develops in tens to hundreds of milliseconds in the absence of misligation [78]. This finding presents a challenge to theoretical models that account for contact order [81], as well as hydrophobic and electrostatic forces [82].

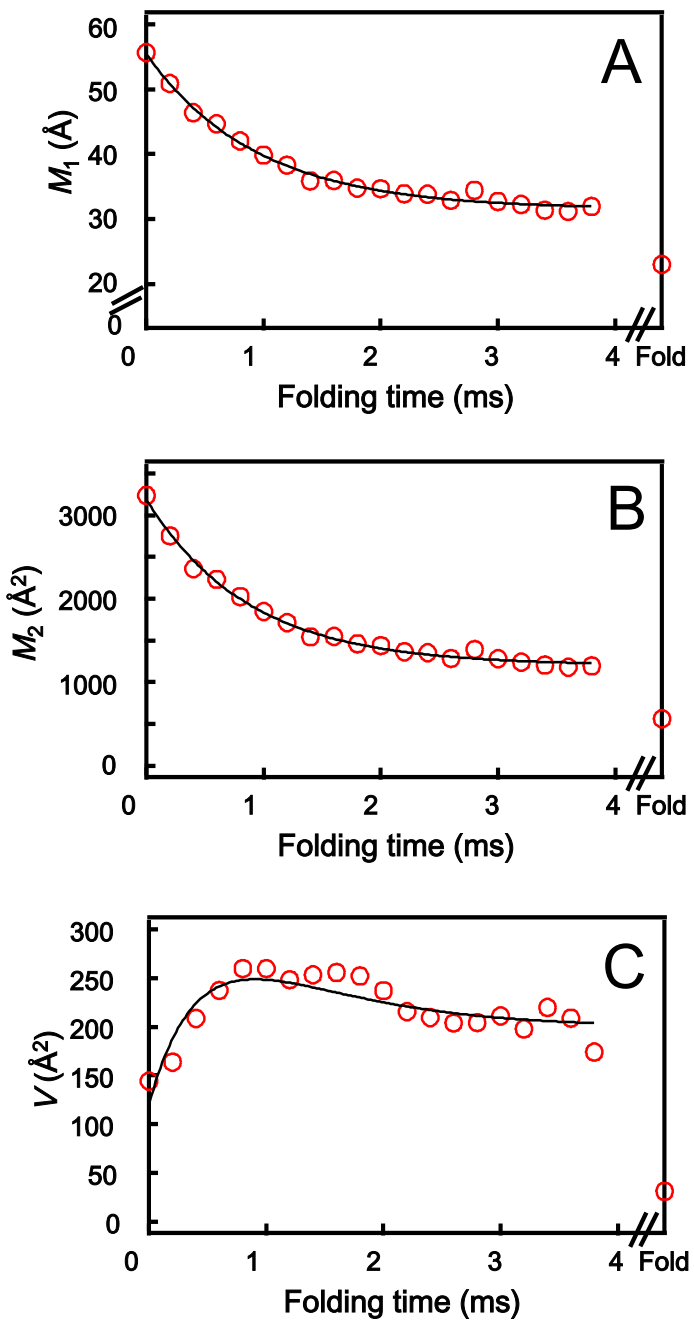

Figure 7.3: Analysis of the P(r_{DA}) fits of the folding kinetics of Dns110-cytochrome c_{552}. (A) Mean distance between Dns110 and the heme (M_1) as a function of folding time. The solid line is a monoexponential fit with a rate constant of 1060 s^{-1}. This fit extrapolates to 31.5 Å, in contrast to M_1 for the folded protein of 23 Å. (B) Time course of the second moment (M_2). The solid line is a monoexponential fit with a rate constant of 1140 s^{-1}. This fit extrapolates to 1200 Å^2, in contrast to M_2 for the folded protein of 560 Å^2. (C) Time course of the variance (V). The solid line is a double exponential fit with rate constants 2070 and 1100 s^{-1}. This fit extrapolates to 200 Å^2, in contrast to V for the folded protein of 31 Å^2.

7.3 Conclusions

Microfluidic mixing by a continuous flow mixer with a 150-μs dead time has enabled us to resolve the fast formation ($\sim$ 1 ms) of the on-pathway intermediate of cytochrome c_{552}. We determined by time-resolved FRET that the intermediate is a heterogeneous ensemble of states with Dns110-heme distances ranging from 15 to $\geq$ 50 Å. This supports the proposal from energy landscape theory that intermediates are a build-up of states due to a kinetic barrier [7, 8], not a discrete structure along a single folding pathway. Our finding likely extends to intermediates in other proteins, emphasizing the importance of using techniques that allow for characterization of heterogeneous ensembles [80].

7.4 Acknowledgments

This work was carried out in collaboration with Seiji Yamada, a visiting scientist from Sony Corporation. We would also like to acknowledge William Ford for his contribution in writing code to aid in the analysis of mixer data.

Chapter 8

Protein Folding Summary

We have investigated the denaturant-induced unfolding of thermostable cytochrome c_{552} by time-resolved fluorescence energy transfer (FRET) between the heme and each of seven site-specific fluorescent probes. The protein exhibits two-state or three-state equilibrium unfolding transitions, depending on the protein region monitored. Interestingly, the observed unfolding intermediate contains native contacts in the unique structural features that differentiate cytochrome c_{552} from other class-I *c*-type cytochromes. This suggests that the extra structural features provide additional stabilization for this thermostable protein.

We then resolved the formation of cytochrome c_{552}'s refolding intermediate by coupling time-resolved FRET to a continuous flow microfluidic mixer. Fitting the rates of energy transfer between a dansyl fluorophore, attached to single-cysteine variants (residue 110 for cytochrome c_{552}), and the cytochrome heme provides us with time-dependent distance distributions, a measure of the reaction progress. The intermediate is comprised of a heterogeneous ensemble of compact and extended polypeptides, an interesting finding that likely extends to other proteins.

We have also characterized the fast folding of cytochrome cb_{562} by time-resolved FRET and microfluidic mixing. Cytochrome cb_{562} belongs to an interesting family of four-helix bundle cytochromes that have nearly identical structures, but widely divergent folding pathways. This rare characteristic creates a unique opportunity to study the relationship between amino acid sequence and folding pathways in proteins with similar structural topologies. The folding rates for two variants that probe folding between the C-terminal heme and either the first helix or the third helix were essentially identical. This suggests that cytochrome c-b_{562} folds by a two-state mechanism in which the entire protein folds cooperatively and rapidly over several milliseconds.

In addition, we have investigated intrachain contact dynamics in unfolded cytochrome cb_{562} by monitoring electron transfer, which occurs as the heme collides with a ruthenium photosensitizer covalently bound to residues along the polypeptide. Intrachain diffusion for chemically denatured proteins proceeds on the microsecond timescale with an upper limit of 0.1 μs. The rate constants exhibit a power-law dependence on the number of peptide bonds between the heme and Ru complex. Adherence of our data to a slope of -1.5, consistent with theoretical models for ideal polymers, demonstrates that cytochrome cb_{562} is minimally frustrated. Importantly, as predicted for a funneled energy landscape, our work shows that making tertiary contacts can enhance the formation of partially folded structures.

Chapter 9

Electron Tunneling Kinetics in Cytochrome cb_{562}

9.1 Introduction

In this chapter, we probe electron tunneling rates in cytochrome cb_{562}. We are interested in improving our fundamental understanding of electron tunneling pathways in proteins. The acquired knowledge can be applied to the design of protein-based electronic devices [83]. For example, to take advantage of the excellent electrocatalytic properties of enzymes in fuel cells, the electronic coupling between the protein's active site and the electrode surface must be optimized [84].

9.1.1 Long-Range Electron Transfer in Proteins

Electron transfer in proteins is a nonadiabatic process that occurs over distances as far as 30 Å through quantum mechanical tunneling. The key relationship that

describes electron transfer processes is the Marcus equation [85]:

$$k_{\text{ET}} = \sqrt{\frac{4\pi^3}{h^2 \lambda RT}} \times H_{AB}^2 \times \exp\left\{-\frac{(\Delta G^\circ + \lambda)^2}{4\lambda RT}\right\} \tag{9.1}$$

where H_{AB} is the electronic coupling matrix element, $-\Delta G^\circ$ is the reaction driving force, and λ is the reorganization energy. H_{AB} exhibits an exponential dependence on distance, represented by β, that becomes particularly relevant in long-range reactions [86].

The exponential decay constant β, which depends on the medium between the two redox-active sites, was probed in activationless systems of different media, including proteins [87, 32]. Two theories are used to explain the β values for different materials, the superexchange model [88] and the uniform barrier model [89]. The tunneling timetable in Figure 9.1 illustrates the distance dependence of electron transfer rates for various proteins. Protein tunneling times are scattered around a β of 1.1 Å^{-1}, most similar to saturated hydrocarbons (1.0 Å^{-1}) and toluene (1.18–1.28 Å^{-1}). The scatter in the data for the α-helical proteins suggests that proteins are not uniform media, which is in fact the case.

Beratan, Onuchic, and co-workers derived a superexchange model for electron transfer in proteins that accounts for the structural complexity of proteins [90, 91, 92]. In this tunneling-pathway model, coupling pathways are broken down into covalent bonds, hydrogen bonds, through-space jumps, and van der Waals interactions, which are combined to create a pathway-specific β value.

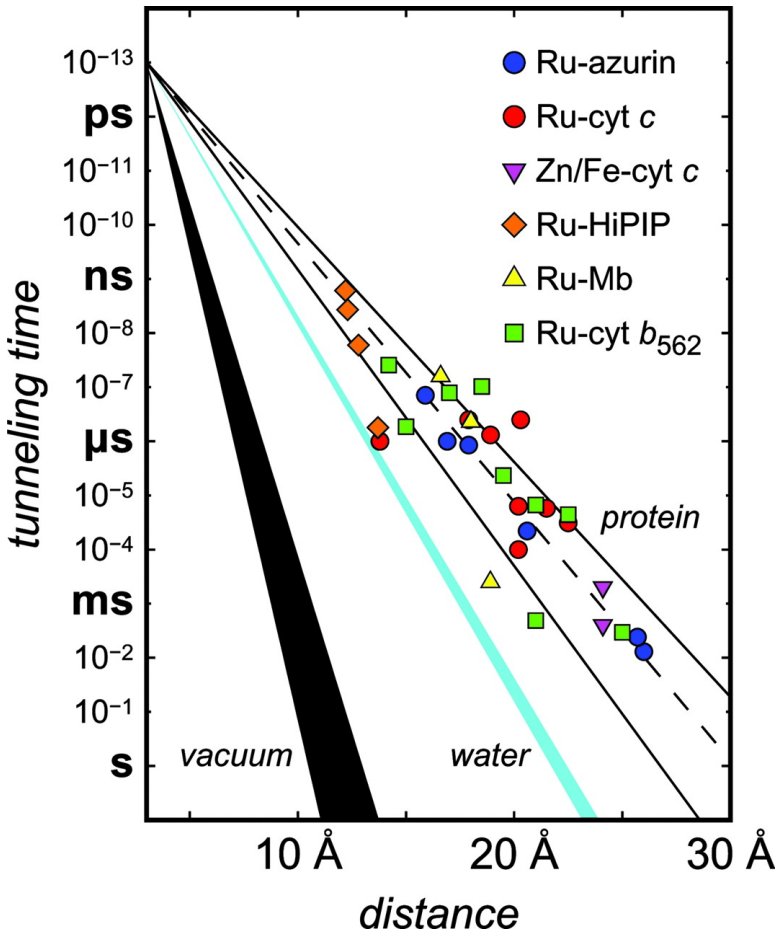

Figure 9.1: Tunneling timetable for proteins. The solid lines illustrate tunneling-pathway predictions for coupling along β-strands (β= 1.0 Å^{-1}) and α-helices (β= 1.3 Å^{-1}). Ru-modified cytochrome b_{562} time constants are represented by green squares. Duplicated with permission from Gray et al. [32].

9.1.2 Electron Tunneling in Cytochromes b_{562} and cb_{562}

The Gray Group previously characterized the electron tunneling rates through the α-helices of Ru-modified cytochrome b_{562} (Figure 9.1) [93, 94]. The data are not well described by a single β value. Prytkova, Kurnikov, and Beratan modeled the electron tunneling pathways in cytochrome cb_{562} and were able to fit seven of the nine variants to a β of 1.3 Å^{-1} [95]. They found that, in the seven variants, electrons tunneled through the heme edge via multiple accessible pathways. The averaging of multiple coupling pathways led to the insensitivity of rates to structure.

In contrast, tunneling rates were slower for the other two variants (His12 and His73) where electrons tunneled to the Fe along a single, dominant pathway through an axial ligand. Thus, dominant pathways are especially sensitive to the protein structural features along the pathway. By this line of reasoning, structural changes to the heme that cause changes in the tunneling mechanism should also alter tunneling rates, even though the distance is conserved.

The protein was modified to cytochrome cb_{562} by introducing covalent linkages from the heme porphyrin to the protein backbone, as depicted in Figure 3.13 and described in Chapter 3. In this modified protein, the heme is coordinated more rigidly to the protein than in the *b*-type cytochrome, where the heme is bound only by axial ligations to a methionine and a histidine residue. We investigate the effects of these changes to the protein structure, i.e., increased rigidity and potential new equatorial tunneling pathways through the covalent linkages, on electron transfer rates in a series of Ru-modified variants (Figure 9.2).

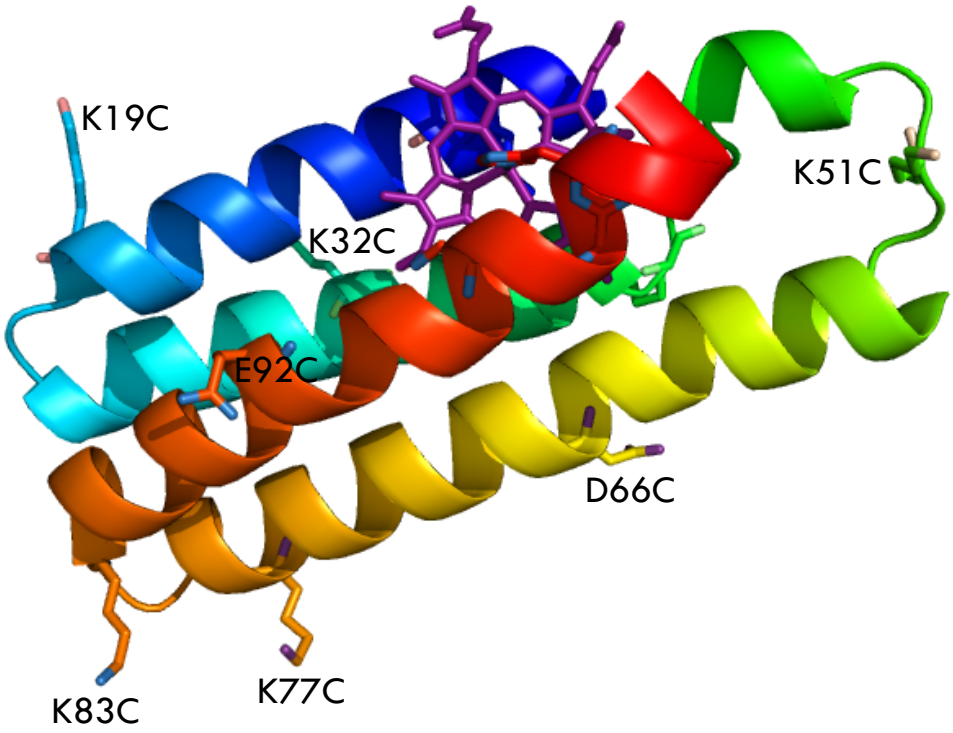

Figure 9.2: We measure Ru–heme electron tunneling rates in cytochrome cb_{562}. $[Ru(bpy)_2(IA\text{-}phen)]^{2+}$ is coupled covalently to surface-exposed cysteines, distributed throughout the protein in single-mutation variants. We have confirmed by circular dichroism and UV-visible spectroscopy that the cysteine mutations and Ru complex do not alter the structure of the protein (section 4.3).

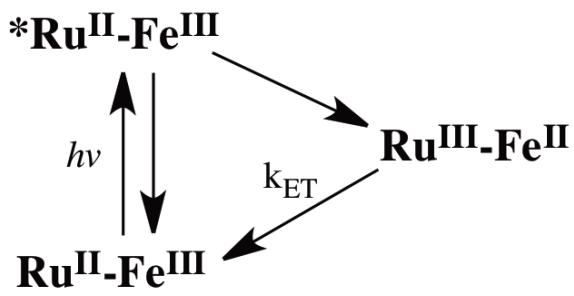

Figure 9.3: Photochemistry of the electron transfer reaction. The back electron transfer rate constant is designated k_{ET}.

9.1.3 Photochemical Triggers

We trigger electron transfer photochemically (Figure 9.3). The Ru label is photoexcited at 480 nm, forming the highly reducing species $*Ru^{II}$. This excited state species can reduce the heme Fe^{III}. For electron transfer to occur, it must be faster than the rate of $*Ru^{II}$ relaxation, or luminescence. We observe the forward ($*Ru^{II}$-Fe^{III}) and back (Fe^{II}-Ru^{III}) electron transfer steps and report the back electron transfer rate constants (k_{ET}).

In studies of cytochrome b_{562}, single-histidine variants were labeled with $[Ru(bpy)_2(imidazole)(His\text{-}imidazole)]^{2+}$. We present data for cytochrome cb_{562} with a different photosensitizer ($[Ru(bpy)_2(IA\text{-}phen)]^{2+}$, Figure 1.7) that is bound covalently to cysteine residues by a flexible linker. The advantages of this new Ru photosensitizer are that it has a longer lifetime (800 ns versus 70 ns), cysteine variants may now be labeled, and the labeling reaction is simpler. A major disadvantage is the uncertainty in the Ru–Fe distance introduced by the flexible linker. A collaboration is underway with Tatiana Prytkova to model the tunneling pathways and the corresponding distance.

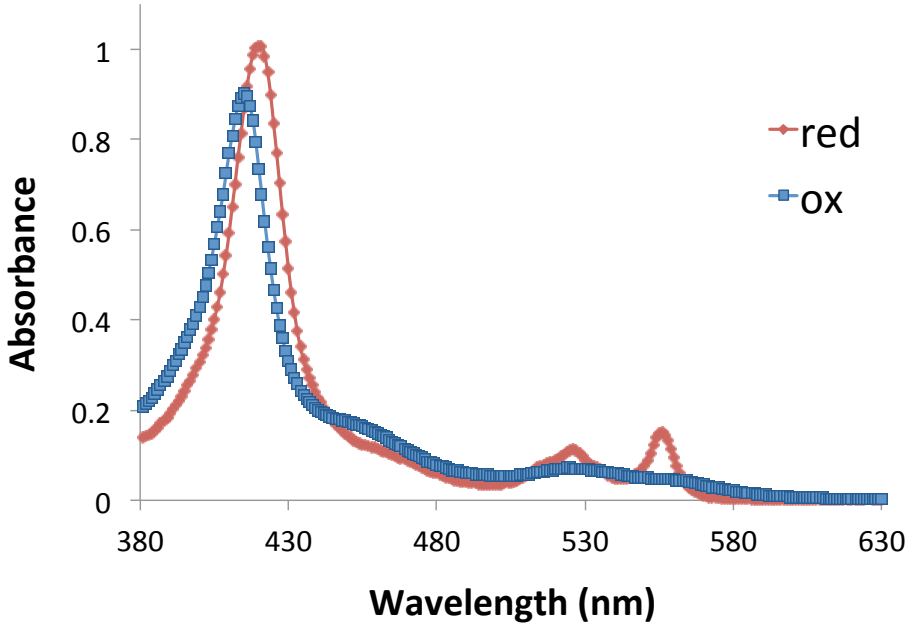

Figure 9.4: Absorption spectrum of oxidized and reduced cytochrome cb_{562}. Upon reduction of Ru-ferricytochrome cb_{562} at pH 7, the Soret maximum shifts from 415 nm to 420 nm.

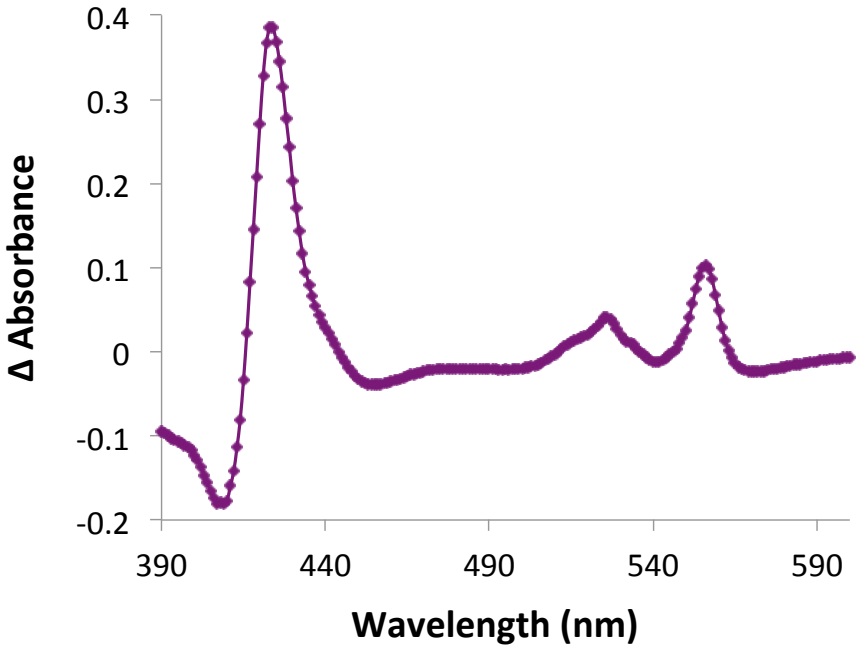

Figure 9.5: The difference spectrum between oxidized and reduced cytochrome cb_{562} exhibits maximal absorbance changes at 409 and 424 nm. We monitor transient absorption kinetics at these wavelengths.

9.2 Difference Absorption Spectrum

To measure electron tunneling rates, the transient species Fe^{II} is monitored at wavelengths that show maximal absorbance changes. The Soret peak of cytochrome cb_{562} red shifts upon reduction (Figure 9.4). The difference spectrum of the oxidized and reduced protein indicates that the largest absorbance changes following reduction occur at 409 and 424 nm (Figure 9.5).

9.3 Kinetic Data

The quenching of the $*Ru^{II}$ luminescence ($k = 1.30 \pm 0.05 \times 10^6\ s^{-1}$ at pH 7) and the corresponding formation of Fe^{II} are indicative of electron transfer. Luminescence decays for the variants are depicted in Figure 9.6, and the rate constants from exponential fits are summarized in Table 9.1. Luminescence decay rates are consistent with exponential fits of the forward electron transfer step, measured by transient absorption.

Transient absorption signals for the Ru-variants are presented in Figures 9.7–9.12. The data fit to biexponential functions where the rate constants correspond to the forward and back electron transfer steps. (The large difference in forward and back rates for Ru66 complicated fits to a biexponential function; instead, the back electron transfer step was fit to a monoexponential function.) The back electron transfer rate constants are summarized in Table 9.1. Electron transfer is not observed for Ru77 and Ru83; the signal decay fits to a single exponential that

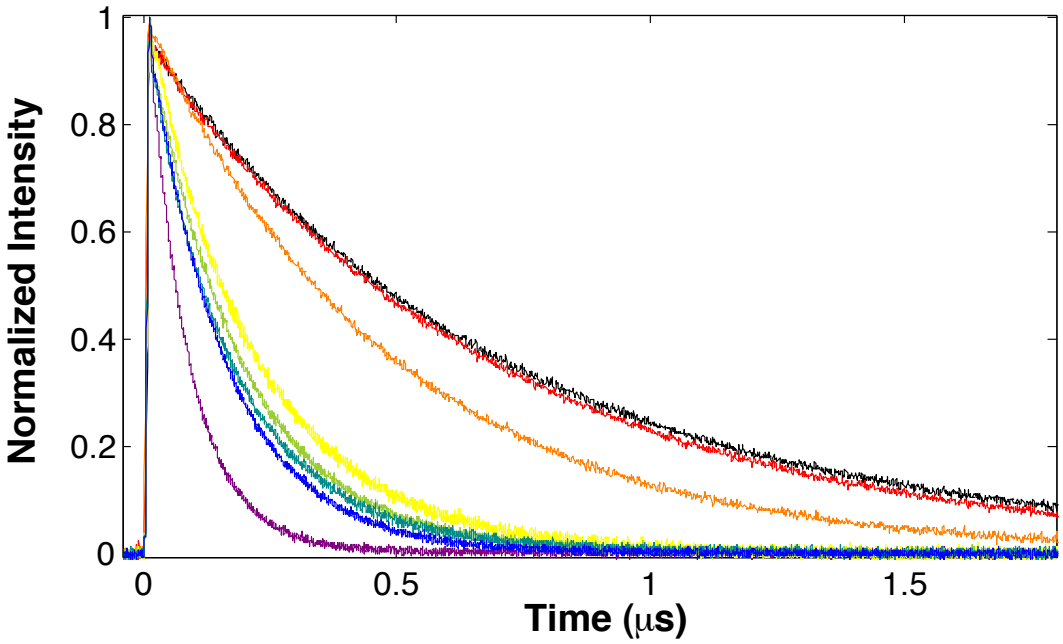

Figure 9.6: Luminescence decays of Ru-modified cytochrome cb_{562} variants and free label. With increasing decay times—Ru32 (violet), Ru92 (blue), Ru51 (teal), Ru19 (green), Ru66 (yellow), Ru77 (orange), Ru83 (red), and label (black).

corresponds to the luminescence rate. Electron transfer is likely too slow at these distances to be competitive with the relaxation of $*Ru^{II}$ by luminescence.

At 424 and 409 nm, $*Ru^{II}$ has a smaller absorbance than Ru^{II} and causes a bleach at time zero. As Fe^{II} has a larger absorbance than Fe^{III} at 424 nm, formation of Fe^{II} is observed as a growth in intensity, so that the signal crosses the baseline in the first kinetic process. In contrast, at 409 nm, Fe^{II} has a negative change in absorbance and is observed following the decay of $*Ru^{II}$ in the first kinetic process. During the second kinetic process, the curve decays to the baseline as the reactant (Ru^{II}-Fe^{III}) is recreated.

Transient ruthenium species can be monitored at 460 nm (Figure 9.11), where the difference in the absorbance of Fe^{III} and Fe^{II} is small (Figure 9.5). We fit

Table 9.1: Luminescence and electron tunneling back rate constants in Ru-modified cytochrome cb_{562} variants.

Variant	Distance (Å)†	Luminescence (s^{-1})	k_{ET} (s^{-1})/424 nm	k_{ET} (s^{-1})/409 nm
Ru32	21.3	1.15 ($\pm$ 0.05) x 10^7	4.1 ($\pm$ 0.5) x 10^6	2.8 ($\pm$ 0.3) x 10^6
Ru92	21.3	6.40 ($\pm$ 0.07) x 10^6	3.7 ($\pm$ 0.1) x 10^5	3.7 ($\pm$ 0.1) x 10^5
Ru51	22.2	5.7 ($\pm$ 0.1) x 10^6	2.7 ($\pm$ 0.3) x 10^5	3.0 ($\pm$ 0.3) x 10^5
Ru19	24.8	5.35 ($\pm$ 0.07) x 10^6	2.8 ($\pm$ 0.3) x 10^5	2.6 ($\pm$ 0.1) x 10^5
Ru66	19.9	4.6 ($\pm$ 0.3) x 10^6	2.3 ($\pm$ 0.4) x 10^3	2.3 ($\pm$ 0.4) x 10^3
Ru77	29.4	2.0 ($\pm$ 0.1) x 10^6	not observed	not observed
Ru83	33.0	1.6 ($\pm$ 0.2) x 10^6	not observed	not observed

†C_γ–Fe distance measured in crystal structure (PDB 2BC5) + 5 Å (approximation).

the forward ($*Ru^{II} \rightarrow Ru^{III}$) and back ($Ru^{III} \rightarrow Ru^{II}$) electron transfer kinetics with biexponential functions. The decay of $*Ru^{II}$ was also monitored at 370 nm (Figure 9.11).

Based on the distance estimates, the rate constants for Ru32 and Ru66 noticeably diverge from the distance dependence. Tunneling to Ru32 is an order of magnitude faster than to Ru92, although it is estimated to be the same distance from the heme. Ru66 is several angstroms closer to the heme, but tunneling is two orders of magnitude slower. As discussed, there is uncertainty in the distance measurement; the estimate is based on the cysteine C_γ–Fe distance, however, the Ru linker could face toward or away from the heme. After obtaining distances from the calculations, we will determine a β for tunneling in cytochrome cb_{562}, identify the outliers, and compare the tunneling pathways.

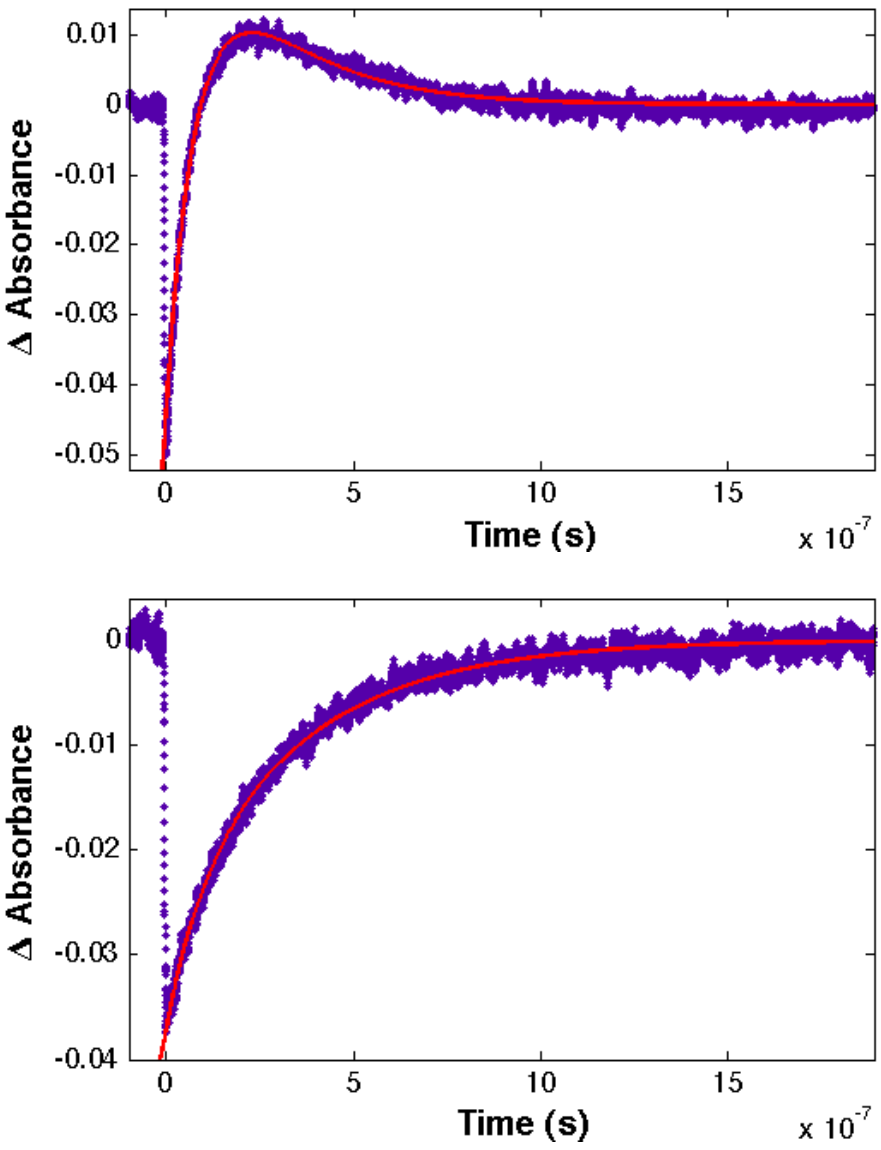

Figure 9.7: Transient absorption spectra of Ru32 at 424 nm (upper) and 409 nm (lower).

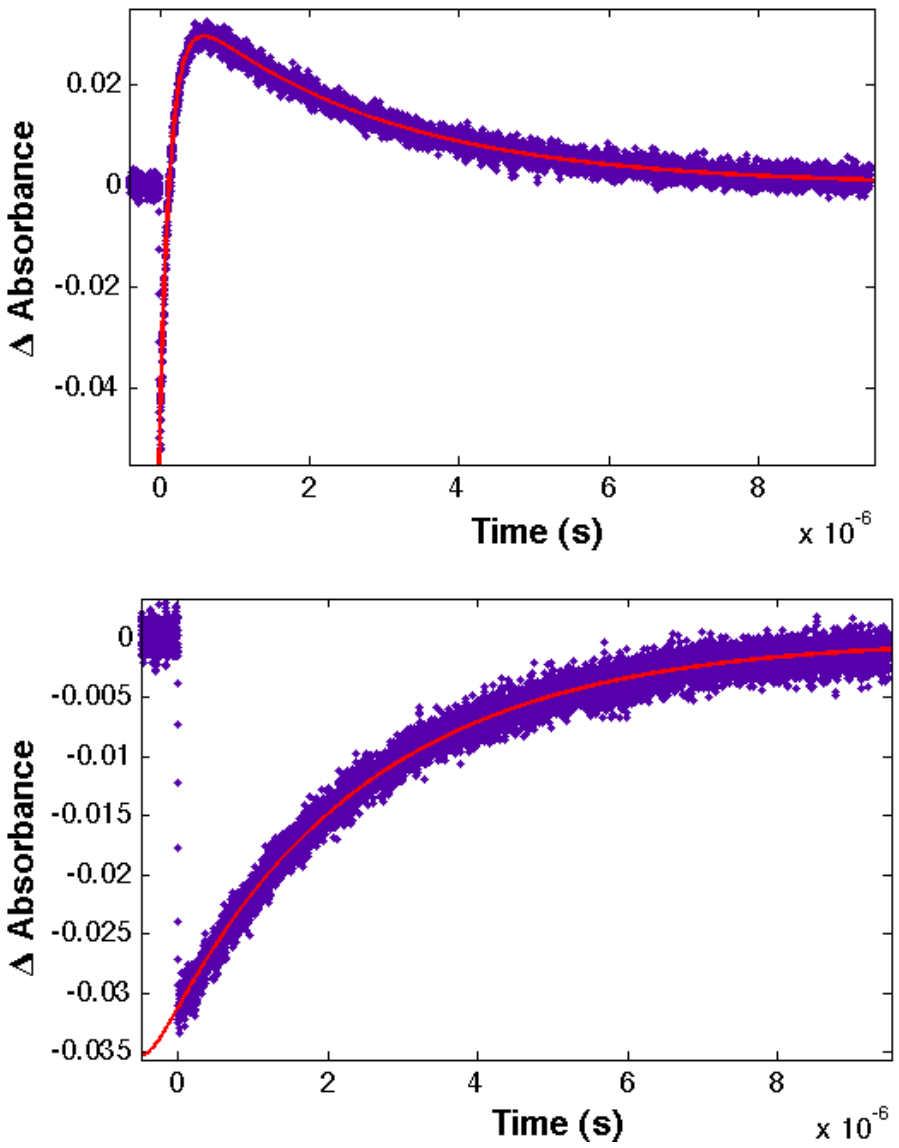

Figure 9.8: Transient absorption spectra of Ru92 at 424 nm (upper) and 409 nm (lower).

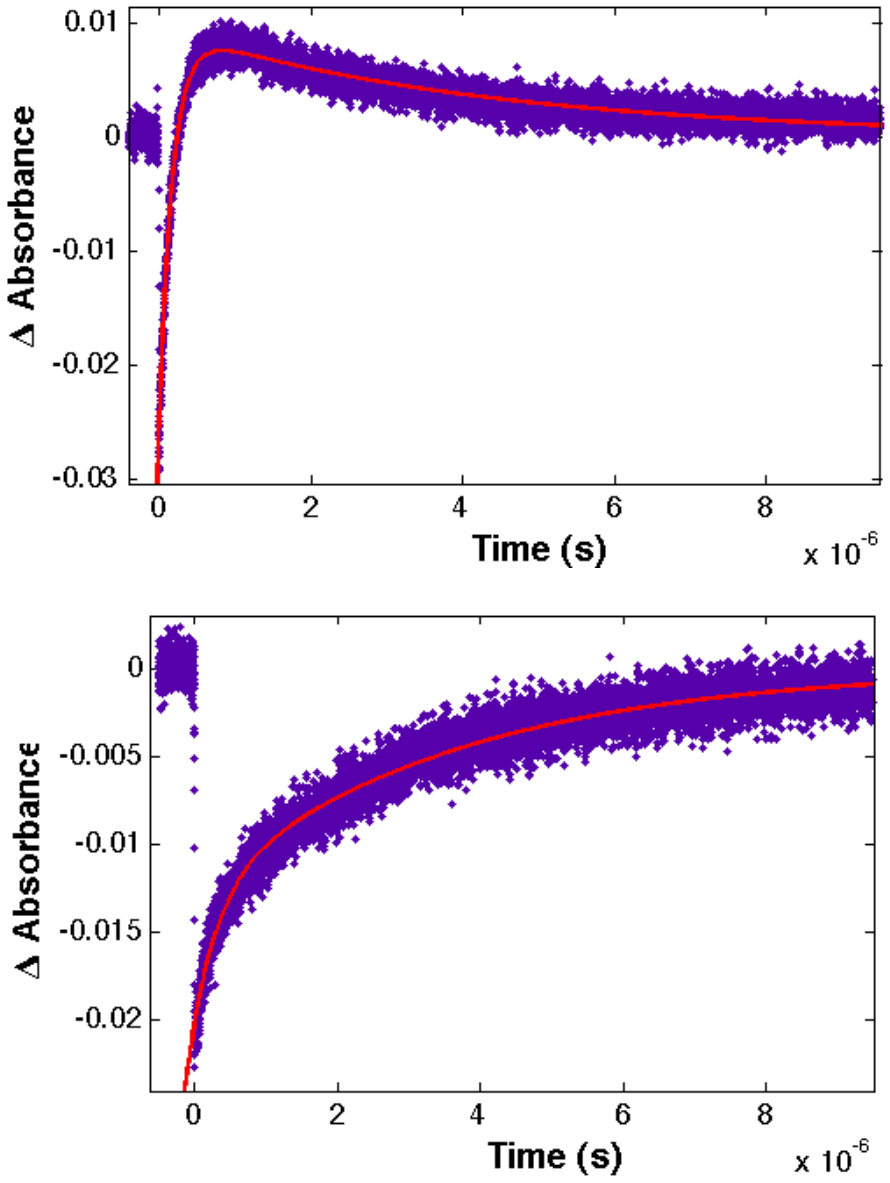

Figure 9.9: Transient absorption spectra of Ru51 at 424 nm (upper) and 409 nm (lower).

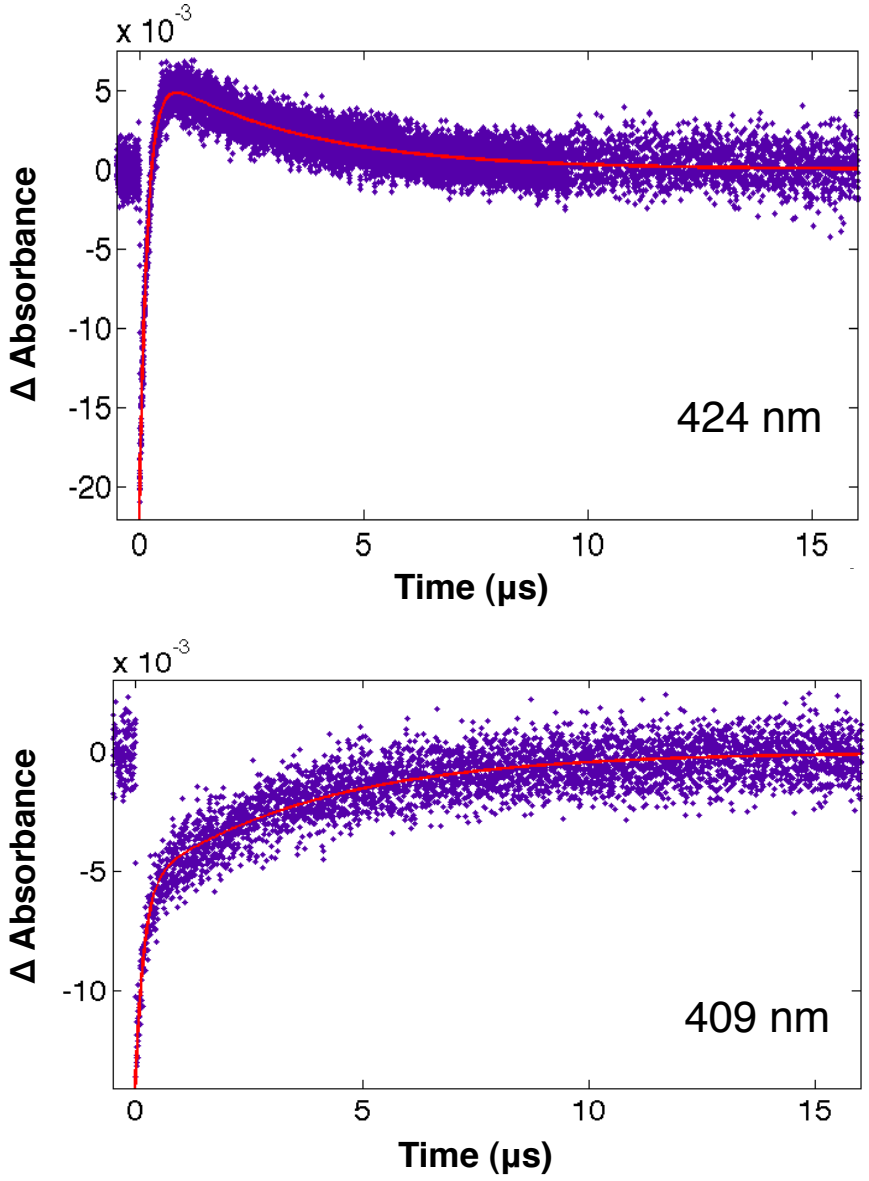

Figure 9.10: Transient absorption spectra of Ru19 at 424 nm (upper) and 409 nm (lower).

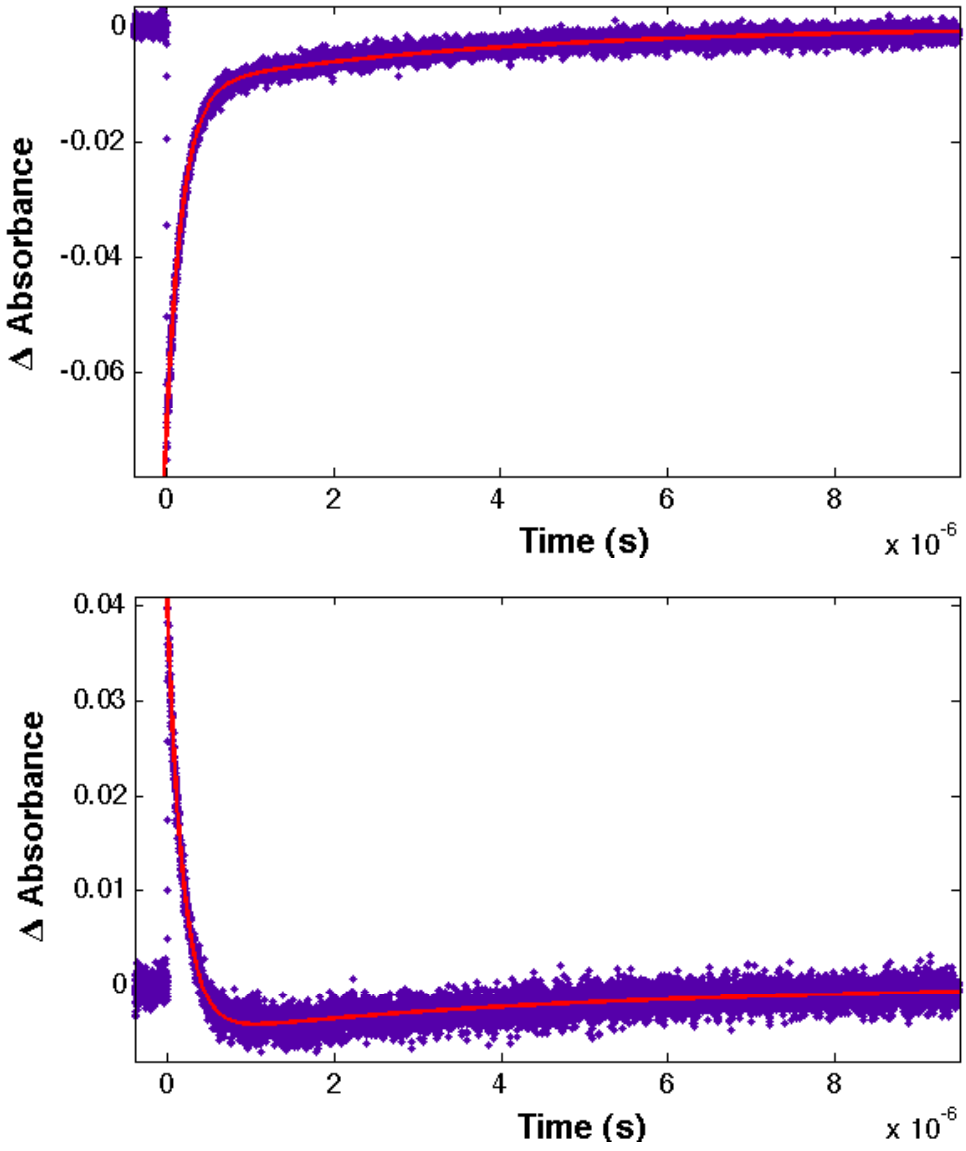

Figure 9.11: Representative transient absorption spectra at 460 nm (upper) and 370 nm (lower), Ru19.

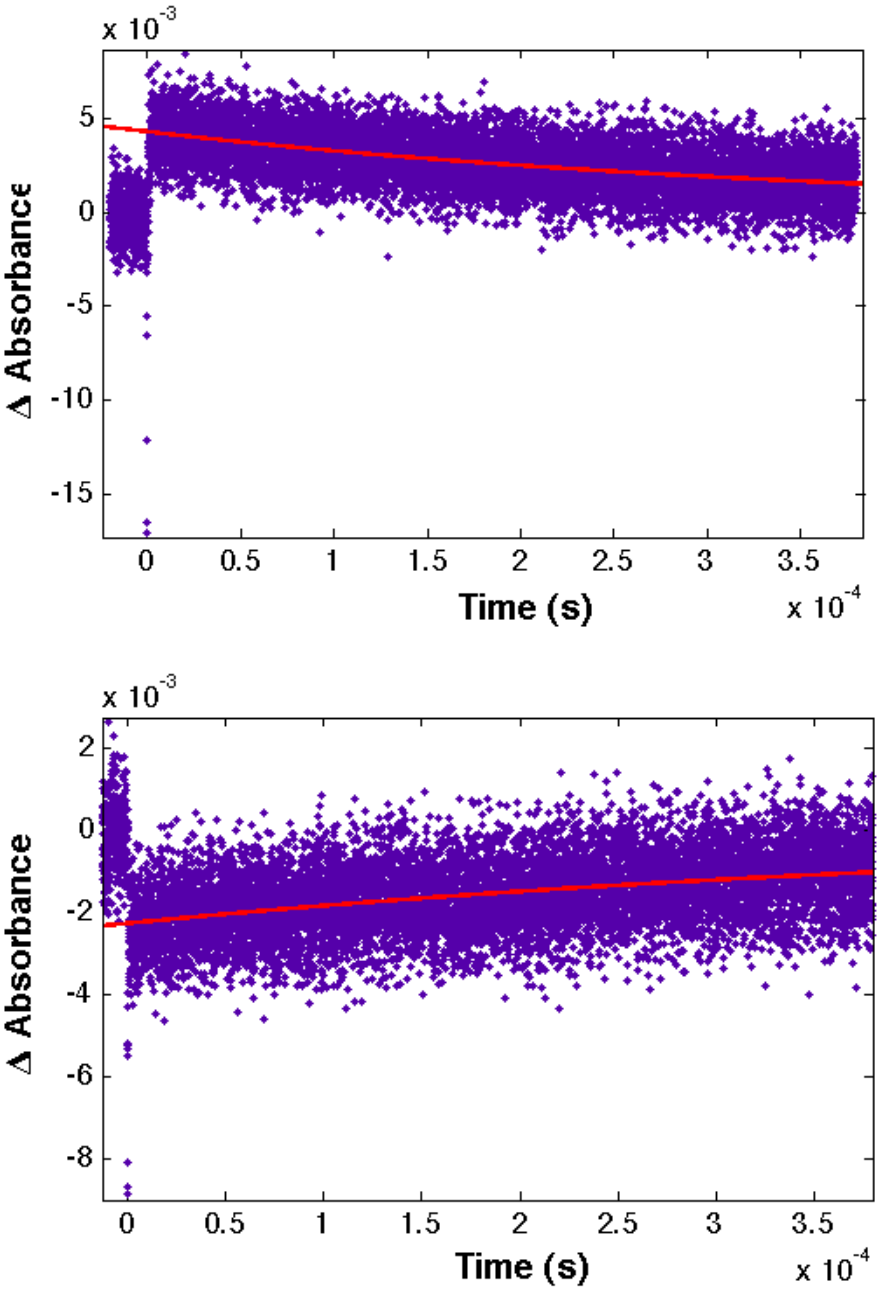

Figure 9.12: Transient absorption spectra of Ru66 at 424 nm (upper) and 409 nm (lower).

Experiments are underway to compare electron tunneling rates in cytochromes cb_{562} and b_{562} in corresponding variants—K19C, K19H, E92C, and E92H. Preliminary results indicate no change in the tunneling rates in Ru-Cys19 with the structural changes to the heme. Tunneling pathways are likely unchanged for Ru-Cys19 or Ru-His19 variants on helix 1 as the covalent porphyrin linkages are on opposite edge of the heme. In contrast, tunneling between residue 92 on helix 4, the covalently linked helix, and the heme could involve the engineered covalent porphyrin linkages. A significant change in the rate would suggest changes in the tunneling mechanism, from multiple pathways to a single dominant pathway.

9.4 Conclusions

With approximations of the Ru–Fe distance, we identify two outliers from the exponential dependence of rate constants on tunneling distance—position 32 on helix 2, which is very fast, and position 66 on helix 3, which is significantly slower than expected. Calculations will provide more accurate distances and insight into tunneling pathways. Additional variants will allow for direct comparisons of tunneling pathways in cytochromes cb_{562} and b_{562}.

9.5 Acknowledgments

I would like to thank Heather Williamson for experimental advice when I started this project and helpful discussions throughout our time working together. I would also like to acknowledge Tatiana Prytkova (Chapman University), a collaborator who is tackling this challenging problem with computational methods.

References

[1] Chiti, F.; Dobson, C. *Annu. Rev. Biochem.*, **2006**, *75*, 333–366.

[2] Steenhuysen, J. Roche Alzheimer's drug picked for major test, May 2012.

[3] Dill, K. A.; Ozkan, S. B.; Shell, M. S.; Weikl, T. R. *Annu. Rev. Biophys.*, **2008**, *37*, 289–316.

[4] Abrol, R.; Griffith, A. R.; Bray, J. K.; Goddard III, W. A. in *Membrane Protein Structure and Dynamics*, Vaidehi, N.; Klein-Seetharaman, J., Eds., volume 914 of *Methods in Molecular Biology*, pp 237–254. Humana Press, 2012 edition, 2012.

[5] Bartlett, A. I.; Radford, S. E. *Nat. Struct. Mol. Biol.*, **2009**, *16*(6), 582–588.

[6] Bowman, G. R.; Voelz, V. A.; Pande, V. S. *Curr. Opin. Struct. Biol.*, **2011**, *21*(1), 4–11.

[7] Wolynes, P. G.; Onuchic, J. N.; Thirumalai, D. *Science*, **1995**, *267*(5204), 1619–1620.

[8] Onuchic, J. N.; Luthey-Schulten, Z.; Wolynes, P. G. *Annu. Rev. Phys. Chem.*, **1997**, *48*, 545–600.

[9] Oliveberg, M.; Wolynes, P. G. *Q. Rev. Biophys.*, **2005**, *38*(3), 245–288.

[10] Wolynes, P. G.; Eaton, W. A.; Fersht, A. R. *Proc. Natl. Acad. Sci. USA*, **2012**, *109*(44), 17770–17771.

[11] Bryngelson, J. D.; Wolynes, P. G. *Proc. Natl. Acad. Sci. USA*, **1987**, *84*(21), 7524–7528.

[12] Plaxco, K. W.; Simons, K. T.; Baker, D. *J. Mol. Biol.*, **1998**, *277*(4), 985–994.

[13] Ivankov, D. N.; Finkelstein, A. V. *Proc. Natl. Acad. Sci. USA*, **2004**, *101*(24), 8942–8944.

[14] Hagen, S. J.; Hofrichter, J.; Szabo, A.; Eaton, W. A. *Proc. Natl. Acad. Sci. USA*, **1996**, *93*(21), 11615–11617.

[15] Chang, I. J.; Lee, J. C.; Winkler, J. R.; Gray, H. B. *Proc. Natl. Acad. Sci. USA*, **2003**, *100*(7), 3838–3840.

[16] Kubelka, J.; Hofrichter, J.; Eaton, W. A. *Curr. Opinion Struct. Biol.*, **2004**, *14*(1), 76–88.

[17] Lapidus, L. J. *Curr. Opin. Struct. Biol.*, **2013**, *23*(1), 30–35.

[18] Lee, J. C.; Gray, H. B.; Winkler, J. R. *J. Am. Chem. Soc.*, **2005**, *127*(47), 16388–16389.

[19] Lee, J. C.; Lai, B. T.; Kozak, J. J.; Gray, H. B.; Winkler, J. R. *J. Phys. Chem. B*, **2007**, *111*(8), 2107–2112.

[20] Förster, T. *Annalen Der Physik*, **1948**, *437*(1), 55–75.

[21] Pletneva, E. V.; Gray, H. B.; Winkler, J. R. *J. Mol. Biol.*, **2005**, *345*(4), 855–867.

[22] Myers, J. K.; Pace, C. N.; Scholtz, J. M. *Protein Sci.*, **1995**, *4*(10), 2138–2148.

[23] Möglich, A.; Krieger, F.; Kiefhaber, T. *J. Mol. Biol.*, **2005**, *345*(1), 153–162.

[24] Brissette, P.; Ballou, D. P.; Massey, V. *Anal. Biochem.*, **1989**, *181*(2), 234–238.

[25] Shastry, M. C. R.; Sauder, J. M.; Roder, H. *Acc. Chem. Res.*, **1998**, *31*(11), 717–725.

[26] Shastry, M. C. R.; Luck, S. D.; Roder, H. *Biophys. J.*, **1998**, *74*(5), 2714–2721.

[27] Roder, H.; Maki, K.; Cheng, H.; Ramachandra Shastry, M. C. *Methods*, **2004**, *34*(1), 15–27.

[28] Kimura, T.; Lee, J. C.; Gray, H. B.; Winkler, J. R. *Proc. Natl. Acad. Sci. USA*, **2007**, *104*(1), 117–122.

[29] Kimura, T.; Lee, J. C.; Gray, H. B.; Winkler, J. R. *Proc. Natl. Acad. Sci. USA*, **2009**, *106*(19), 7834–7839.

[30] Winkler, J.; Gray, H. *Chem. Rev.*, **1992**, *92*, 369–379.

[31] Ener, M. E.; Lee, Y.-T.; Winkler, J. R.; Gray, H. B.; Cheruzel, L. *Proc. Natl. Acad. Sci. USA*, **2010**, *107*(44), 18783–18786.

[32] Gray, H. B.; Winkler, J. R. *Proc. Natl. Acad. Sci. USA*, **2005**, *102*(10), 3534–3539.

[33] Wittung-Stafshede, P.; Lee, J. C.; Winkler, J. R.; Gray, H. B. *Proc. Natl. Acad. Sci. USA*, **1999**, *96*(12), 6587–6590.

[34] Lee, J. C.; Gray, H. B.; Winkler, J. R. *Proc. Natl. Acad. Sci. USA*, **2001**, *98*(14), 7760–7764.

[35] Lee, J. C.; Engman, K. C.; Tezcan, F. A.; Gray, H. B.; Winkler, J. R. *Proc. Natl. Acad. Sci. USA*, **2002**, *99*(23), 14778–14782.

[36] Faraone-Mennella, J.; Gray, H. B.; Winkler, J. R. *Proc. Natl. Acad. Sci. USA*, **2005**, *102*(18), 6315–6319.

[37] Faraone-Mennella, J.; Tezcan, F. A.; Gray, H. B.; Winkler, J. R. *Biochemistry*, **2006**, *45*(35), 10504–10511.

[38] Pletneva, E. V.; Zhao, Z.; Kimura, T.; Petrova, K. V.; Gray, H. B.; Winkler, J. R. *J. Inorg. Biochem.*, **2007**, *101*(11-12), 1768–1775.

[39] Gray, H. B.; Winkler, J. R.; Kozak, J. J. *Mol. Phys.*, **2011**, *109*(6), 905–916.

[40] Gray, H. B.; Winkler, J. R.; Kozak, J. J. *Mol. Phys.*, **2012**, *110*(7), 419–429.

[41] Lederer, F.; Glatigny, A.; Bethge, P. H.; Bellamy, H. D.; Matthew, F. S. *J. Mol. Biol.*, **1981**, *148*(4), 427–448.

[42] Shibata, N.; Iba, S.; Misaki, S.; Meyer, T. E.; Bartsch, R. G.; Cusanovich, M. A.; Morimoto, Y.; Higuchi, Y.; Yasuoka, N. *J. Mol. Biol.*, **1998**, *284*(3), 751–760.

[43] Bertini, I.; Faraone-Mennella, J.; Gray, H.; Luchinat, C.; Parigi, G.; Winkler, J. *J. Biol. Inorg. Chem.*, **2004**, *9*, 224–230.

[44] Hill, E. E.; Morea, V.; Chothia, C. *J. Mol. Biol.*, **2002**, *322*(1), 205–233.

[45] Garcia, P.; Bruix, M.; Rico, M.; Ciofi-Baffoni, S.; Banci, L.; Ramachandra Shastry, M. C.; Roder, H.; de Lumley Woodyear, T.; Johnson, C. M.; Fersht, A. R.; Barker, P. D. *J. Mol. Biol.*, **2005**, *346*(1), 331–344.

[46] Brodin, J. D.; Medina-Morales, A.; Ni, T.; Salgado, E. N.; Ambroggio, X. I.; Tezcan, F. A. *J. Am. Chem. Soc.*, **2010**, *132*(25), 8610–8617.

[47] Tokita, Y.; Yamada, S.; Luo, W.; Goto, Y.; Bouley-Ford, N.; Nakajima, H.; Watanabe, Y. *Angew. Chem. Int. Ed.*, **2011**, *50*(49), 11663–11666.

[48] Muresanu, L.; Pristovsek, P.; Lohr, F.; Maneg, O.; Mukrasch, M. D.; Ruterjans, H.; Ludwig, B.; Lucke, C. *J. Biol. Chem.*, **2006**, *281*(20), 14503–14513.

[49] Travaglini-Allocatelli, C.; Gianni, S.; Morea, V.; Tramontano, A.; Soulimane, T.; Brunori, M. *J. Biol. Chem.*, **2003**, *278*(42), 41136–41140.

[50] Nozaki, Y. *Methods Enzymol.*, **1972**, *26*, 43–50.

[51] Yamada, S.; Bouley Ford, N. D.; Keller, G. E.; Ford, William C; Gray, H. B.; Winkler, J. R. *Proc. Natl. Acad. Sci. USA*, **2013**, *110*(6), 1606–1610.

[52] Hon-Nami, K.; Oshima, T. *J. Biochem.*, **1977**, *82*(3), 769–776.

[53] Beals, J. M.; Haas, E.; Krausz, S.; Scheraga, H. A. *Biochemistry*, **1991**, *30*(31), 7680–7692.

[54] Beechem, J. M.; Haas, E. *Biophys. J.*, **1989**, *55*(6), 1225–1236.

[55] Lawson, C. L.; Hanson, R. J. *Solving Least Squares Problems*. Prentice-Hall, Englewood Cliffs, NJ, 1974.

[56] Wu, P.; Brand, L. *Anal. Biochem.*, **1994**, *218*(1), 1–13.

[57] Kimura, T.; Lee, J. C.; Gray, H. B.; Winkler, J. R. *Proc. Natl. Acad. Sci. USA*, **2007**, *104*(1), 117–122.

[58] Akiyama, S.; Takahashi, S.; Kimura, T.; Ishimori, K.; Morishima, I.; Nishikawa, Y.; Fujisawa, T. *Proc. Natl. Acad. Sci. USA*, **2002**, *99*(3), 1329–1334.

[59] Pace, N. C.; Shirley, B. A.; Thomson, T. F. *Protein Structure: A Practical Approach*. IRL, Oxford, 1990.

[60] Hagen, S. J.; Hofrichter, J.; Eaton, W. A. *J. Phys. Chem. B*, **1997**, *101*(13), 2352–2365.

[61] Szabo, A.; Schulten, K.; Schulten, Z. *J. Chem. Phys.*, **1980**, *72*(8), 4350–4357.

[62] Thirumalai, D. *J. Phys. Chem. B*, **1999**, *103*(4), 608–610.

[63] Lapidus, L. J.; Eaton, W. A.; Hofrichter, J. *Proc. Natl. Acad. Sci. USA*, **2000**, *97*(13), 7220–7225.

[64] Krieger, F.; Fierz, B.; Bieri, O.; Drewello, M.; Kiefhaber, T. *J. Mol. Biol.*, **2003**, *332*(1), 265–274.

[65] Perl, D.; Jacob, M.; Bánó, M.; Stupák, M.; Antalik, M.; Schmid, F. X. *Biophys. Chem.*, **2002**, *96*, 173–190.

[66] Urie, K. G.; Angulo, D.; Lee, J. C.; Kozak, J. J.; Gray, H. B.; Winkler, J. R. *J. Phys. Chem. B*, **2009**, *113*(2), 522–530.

[67] Fierz, B.; Kiefhaber, T. *J. Am. Chem. Soc.*, **2007**, *129*(3), 672–679.

[68] Buscaglia, M.; Lapidus, L. J.; Eaton, W. A.; Hofrichter, J. *Biophys. J.*, **2006**, *91*(1), 276–288.

[69] Chu, R.; Pei, W.; Takei, J.; Bai, Y. *Biochemistry*, **2002**, *41*(25), 7998–8003.

[70] Zhou, Z.; Huang, Y.; Bai, Y. *J. Mol. Biol.*, **2005**, *352*(4), 757–764.

[71] Wang, T.; Zhou, Z.; Bunagan, M. R.; Du, D.; Bai, Y.; Gai, F. *Protein Sci.*, **2007**, *16*(6), 1176–1183.

[72] Zhou, Y.; Karplus, M. *Nature*, **1999**, *401*(6751), 400–403.

[73] Chahine, J.; Nymeyer, H.; Leite, V. B.; Socci, N. D.; Onuchic, J. N. *Phys. Rev. Lett.*, **2002**, *88*(16), 168101.

[74] Ferreiro, D. U.; Hegler, J. A.; Komives, E. A.; Wolynes, P. G. *Proc. Natl. Acad. Sci. USA*, **2007**, *104*(50), 19819–19824.

[75] Than, M. E.; Hof, P.; Huber, R.; Bourenkov, G. P.; Bartunik, H. D.; Buse, G.; Soulimane, T. *J. Mol. Biol.*, **1997**, *271*(4), 629–644.

[76] Ibrahim, S. M.; Nakajima, H.; Ohta, T.; Ramanathan, K.; Takatani, N.; Naruta, Y.; Watanabe, Y. *Biochemistry*, **2011**, *50*(45), 9826–9835.

[77] Wain, R.; Pertinhez, T. A.; Tomlinson, E. J.; Hong, L.; Dobson, C. M.; Ferguson, S. J.; Smith, L. J. *J. Biol. Chem.*, **2001**, *276*(49), 45813–45817.

[78] Pletneva, E. V.; Gray, H. B.; Winkler, J. R. *Proc. Natl. Acad. Sci. USA*, **2005**, *102*(51), 18397–18402.

[79] Baldwin, R. L. *Folding & Design*, **1996**, *1*(1), R1–8.

[80] Udgaonkar, J. B. *Annu. Rev. Biophys.*, **2008**, *37*, 489–510.

[81] Clementi, C.; Nymeyer, H.; Onuchic, J. N. *J. Mol. Biol.*, **2000**, *298*(5), 937–953.

[82] Weinkam, P.; Pletneva, E. V.; Gray, H. B.; Winkler, J. R.; Wolynes, P. G. *Proc. Natl. Acad. Sci. USA*, **2009**, *106*(6), 1796–1801.

[83] Tokita, Y.; Shimura, J.; Nakajima, H.; Goto, Y.; Watanabe, Y. *J. Am. Chem. Soc.*, **2008**, *130*(15), 5302–5310.

[84] Cracknell, J. A.; Vincent, K. A.; Armstrong, F. A. *Chem. Rev.*, **2008**, *108*(7), 2439–2461.

[85] Marcus, R.; Sutin, N. *Biochim. Biophys. Acta*, **1985**, *811*, 265–322.

[86] Hopfield, J. *Proc. Natl. Acad. Sci. USA*, **1974**, *71*(9), 3640–3644.

[87] Gray, H. B.; Winkler, J. R. *Q. Rev. Biophys.*, **2003**, *36*(3), 341–372.

[88] McConnell, H. M. *J. Chem. Phys.*, **1961**, *35*(2), 508–515.

[89] Moser, C.; Keske, J.; Warncke, K.; Farid, R.; Dutton, P. *Nature*, **1992**, *355*, 796–802.

[90] Beratan, D.; Onuchic, J. *Photosynth. Res.*, **1989**, *22*(3), 173–186.

[91] Beratan, D. N.; Betts, J. N.; Onuchic, J. N. *Science*, **1991**, *252*(5010), 1285–1288.

[92] Onuchic, J. N.; Beratan, D. N.; Winkler, J. R.; Gray, H. B. *Annu. Rev. Biophys. Biomol. Struct.*, **1992**, *21*, 349–377.

[93] Winkler, J. R.; Di Bilio, A. J.; Farrow, N. A.; Richards, J. H.; Gray, H. B. *Pure Appl. Chem.*, **1999**, *71*(9), 1753–1764.

[94] Farrow, N. *Investigation of Electron Transfer in the Alpha-Helical Protein Cytochrome b562*. PhD Thesis, California Institute of Technology, 1999.

[95] Prytkova, T. R.; Kurnikov, I. V.; Beratan, D. N. *Science*, **2007**, *315*(5812), 622–625.